RELATIONSHIPS AMONG CONSTRUCTS OF L2 CHINESE READING AND LANGUAGE BACKGROUND

Wei-Li Hsu

Dissertation Draft

December 16, 2016

Committee Chair:

Thomas Hudson

Committee Members:

James Dean Brown

Richard Day

Richard Schmidt

Haidan Wang

Keywords: L2 reading, Chinese as a second language, reading strategies, heritage language learning

To my mother, Bi-Yun Chen, and father, Tze-Mao Hsu, whose support, encouragement, and constant love have sustained me throughout my life.

ACKNOWLEDGMENTS

This dissertation would not have been possible without the support and encouragement of many individuals. First of all, I thank my dissertation committee for all the insightful comments and constructive suggestions they made during my journey to this degree. I am particularly indebted to Professor Thom Hudson. His remarkable understanding of second-language reading and language assessment is complemented by his approachability and the patience he shows toward all students. I will always be grateful for my good fortune in having had his guidance throughout my doctoral studies. I also thank Professor J.D. Brown for his research methods courses, which provide easy-to-understand instruction in statistical theories applicable to the field of applied linguistics. Professor Richard Day and Professor Dick Schmidt always provided incisive commentary that pushed me to more clearly link my own research to the wider fields of second-language theory and teaching; and I am also grateful to Professor Haidan Wang for aiding my growth and progress as a lecturer in the Department of East Asian Languages and Literatures.

This dissertation benefited immensely from the financial support of a dissertation grant offered by the journal *Language Learning*, and from a TOEFL Small Grant for Doctoral Research in Second or Foreign Language Assessment, which I greatly appreciate. Last but not least, I would like to thank my family in Hawaii, Richard Kawasaki and Milton Kawasaki, for their support and encouragement, and to acknowledge the help and inspiration I received from the rest of my family, friends, colleagues, and students in the many places I have lived and taught throughout the time I worked on this dissertation, and the coursework that came before it. Thank you. 謝謝. Mahalo nui loa.

ABSTRACT

Extensive research has been conducted on the relationships of Chinese-character recognition to reading development; strategic competence to reading comprehension; and home linguistic exposure to heritage language acquisition. However, studies of these relationships have been marked by widely divergent theoretical underpinnings, and their results are not directly comparable. The current study adopts a cognitive and component perspective on reading, and uses structural equation modeling (SEM) to examine character-recognition skills, reading-comprehension skills, strategy use, and language background. Among these four factors, character-recognition skills are held to represent lower-level reading processing; reading comprehension and strategy-use represent higher-level processing; and language background is operationalized as a source of background knowledge.

The present study's 85 participants were divided into four groups representing four language backgrounds: Singaporean Chinese Mother Tongue language learners (Singaporean CMTLLs) (n = 14), Chinese foreign language learners (CFLLs) (n = 19), Mandarin-speaking heritage language learners (Man-HLLs) (n = 38), and Cantonese-speaking heritage language learners (Can-HLLs) (n = 14). A package of eight instruments was administered via an online platform, and included a language-background survey, a multiple-choice grammar subtest, a fill-in-the-blank cloze subtest, a multiple-choice passage-comprehension subtest, a strategy-use survey, an ortho-phonological subtest, an ortho-semantic subtest, and a morpheme-discrimination subtest. The language-background survey was designed to gain a clear understanding of the participants' language backgrounds; the grammar, cloze, and passage-comprehension subtests, to investigate their reading-comprehension ability; the strategy-use survey, to capture the participants' perceived use of six strategy types; and the ortho-

phonological, ortho-semantic, and morpheme subtests, to examine their character-recognition ability.

Prior to examining common constructs of L2 Chinese reading across the participants' four language backgrounds, three profile analyses were conducted to examine the extent of language-background effects on reading comprehension, character-recognition, and strategy use. The results suggested that language background had a stronger effect on reading-comprehension ability ($p = 0.001$, partial $\eta^2 = 0.18$), than on character-recognition ability ($p > 0.05$, partial $\eta^2 = 0.09$) or on strategy use ($p > 0.05$, partial $\eta^2 = 0.05$).

Next, a hypothesized SEM model was examined and modified. All parameter estimates in the revised SEM model were statistically significant at $p < 0.05$, and the 12 variables explained 99% of the variance in reading-comprehension ability. The structural model of the revised SEM model indicated that character recognition had a strong and direct effect on reading comprehension ($r = 0.99$); cognitive-strategy use had a medium and direct effect on reading comprehension ($r = 0.21$); and metacognitive-strategy use had no direct influence on reading comprehension, but yielded a strong and direct influence on cognitive-strategy use ($r = 0.71$), supporting the notion that metacognitive strategies exert an executive function over cognitive ones. However, monitoring – a type of metacognitive strategy – directly and negatively influenced character-recognition ($r = -0.34$). These paths suggest that the participants constantly monitored their character-level processes, and that when they encountered unfamiliar characters, monitoring strategies activated other metacognitive strategies to regulate cognitive strategies, which in turn facilitated comprehension processes to compensate for the interrupted character-recognition processes.

The results have practical implications for Chinese language learning and second language acquisition (SLA) more generally, and examining L2 Chinese reading with SLA theories.

TABLE OF CONTENTS

List of Tables

List of Figures

CHAPTER 1

INTRODUCTION

1.1 Context of the Problem

As the number of Americans enrolling in Chinese language classes increases, more needs to be done to understand how people learn Chinese as a second language (L2). In higher education, reading skills are often regarded as the foundation of L2 development, and when learning how to read Chinese script, Chinese characters are reported as a major challenge to beginning readers (Everson & Ke, 1997; Shen, 2013). Responding to this challenge, researchers have investigated the development of Chinese character acquisition and its role in L2 Chinese reading (Everson, 2011; Ke, 1998; Koda, Zhang, & Yang, 2008; Wang & Leland, 2011; Xiao, 2008). However, focusing specifically on Chinese character acquisition would lead to overlooking the interactive nature of reading acts and that reaching adequate comprehension is more complicated than the sum of all the pieces of decoded information (Cohen & Upton, 2007; Grabe, 2009; Hudson, 2007; Stanovich, 2000; Alexander & Fox, 2013). In addition to an interactive perspective of L2 Chinese reading, I am interested in how well the interactive model of L2 Chinese reading can be applied to various groups of Chinese language learners with different backgrounds. Therefore, the main purpose of this dissertation is to examine constructs of L2 Chinese reading across language backgrounds from an interactive perspective. As well as my general perspectives on L2 Chinese reading, this chapter explains my rationales for choosing particular processes, the definitions of the terms used in this dissertation, and the importance of the study.

1.2 A Cognitive Perspective on L2 Chinese Reading

Usually, reading ability is measured through assessing the degree of comprehension after reading printed materials. Given that reading material can be as short as a word or as long as a novel, cognitive theories operationalize reading ability as mental processes, or micro-processes, dealing with word-level, phrase-level, sentence-level, paragraph-level of literal and inferential understanding (Alderson, 2000). Due to the component view of reading processes, it is important to note that I am aware that the distinction between component skills is artificial and these skills often overlap with each other to varying degrees. Bearing its limitations in mind, an advantage of the component-skill perspective is that it provides an operationalized framework for measuring various reading skills and comprehension levels, so that researchers are able to go beyond observable behaviors and tap into unobservable cognitive mechanisms. In the case of assessing L2 reading ability – in contrast to speaking and writing abilities, which can be examined through language performance – the underlying cognitive mechanism involved in reading acts cannot be directly observed, and the theorized reading components allow scholars to make inferences about reading ability by eliciting behaviors that correspond to the component skills. Although I am aware that cognitive theories have been heavily criticized for neglecting the social aspect of L2 acquisition and relating human minds to computer processors (Block, 2003), this dissertation adopts a cognitive perspective due to the strength of its ability to assess unobservable mental operations through eliciting responses that are not influenced by speaking or writing abilities.

1.3 An Interactive Perspective on L2 Chinese Reading

When approaching texts for the purpose of comprehending them, readers engage in multiple reading processes, and these processes are generally categorized as either higher- or lower-level. Lower-level processes often refer to word-recognition skills, such as identifying

printed words, assigning meaning to them, and building connections between and among these words, their pronunciations, and their meanings (Grabe, 2009; Hudson, 2007). Scholars who focus on lower-level processes have tended to adopt a linear perspective on reading processes: namely, that comprehension is achieved through combining smaller pieces of decoded information from printed words. Higher-level processes of comprehension, on the other hand, can be subdivided into three main types – 1) building literal comprehension, 2) constructing the reader's interpretation of a text, and 3) directing attentional resources to particular lower- and higher-level processes – and scholars focusing on these higher-level processes tend to view reading acts as recursive, interactive, and goal-oriented (Grabe, 2009; Perfetti, 1999). With regard to studies of first language (L1) reading and L2 reading, however, most scholars agree (regardless of their own focus on lower- or higher-level processes) that both lower and higher levels of process are necessary for fluent reading; and that interactions between various processes are necessary (Grabe, 2009; Koda, 2005; Hudson, 2007; Stanovich, 2000). Accordingly, this dissertation adopts an interactive perspective that recognizes the contribution of both lower- and higher-level processes to adequate reading comprehension.

1.4 Rationale of Chosen Reading Processes in the Dissertation

Due to constraints on time and resources, three groups of skills are selected, which can be divided into two broad categories, i.e., lower-level and higher-level processes, with one group falling into the former and two into the latter category. The group of the lower-level process is character-recognition ability, and the two groups of the higher-level processes are comprehension ability and strategy-use.

1.4.1 Lower-level processes. Although most scholars recognize the importance of interaction between levels and processes, word-recognition tends to be the first challenge

experienced by beginning language learners (Hudson, 2007). Given the profound differences between Chinese orthography and alphabetic orthographies, it is perhaps unsurprising that learning Chinese characters is one of the most common challenges reported by beginning Chinese language learners (Everson, 1998, 2011; Zhao, Guo, & Dynia, 2013).

To reflect the challenge relating to character acquisition, many researchers investigate Chinese-character recognition as a lower-level process for people who are bilingual in Chinese and English (for L2 Chinese readers: Everson, 1998; Lü & Koda, 2011; Shen & Ke, 2007; for bilingual Hong Kong children: McBride-Chang & Kail, 2002; Shu, Chen, Anderson, Wu, & Xuan, 2003; Tong & McBride-Chang, 2010). Such studies have shown that the ability to identify and analyze structural features of Chinese characters is positively correlated with Chinese-character recognition (Koda, 2005; Kuo & Anderson, 2008; McBride-Chang, Zhou, Cho, Aram, Levin, & Tolchinsky, 2011).

1.4.2 Higher-level processes. Two of the three main areas of reading comprehension are literal comprehension and readers' interpretations; these two areas can be further divided into several higher-level processes, such as organizing deciphered information from lower-level processes, forming coherent mental representations of texts, and interpreting texts based on background knowledge (Grabe, 2009; Perfetti, 1999). In light of the relationship between background knowledge and reading comprehension (Grabe, 2009; Khalifa & Weir, 2009), it would appear worth investigating whether Chinese L2 learners with extensive prior exposure to their target language would demonstrate better reading test performance due to their broader background knowledge than Chinese foreign language learners do.

In addition to literal comprehension and readers' interpretation, the third area of reading-comprehension processes is strategy use, also known as strategic competence (Grabe, 2009;

Hudson, 2007; Phakiti, 2008; Purpura, 1997; Rupp, Ferne, & Choi, 2006), which enhances language use in specific settings, such as L2 reading tests, under a communicative framework of L2 ability (Bachman & Palmer, 1996; Canale & Swain, 1980; Cohen, 2014). Strategic competence initially interested researchers as a compensatory function for reparing communication breakdowns (Canale & Swain, 1980), but later studies (Phakiti, 2008; Purpura, 1997) found that some strategies were also positively correlated with reading-test performance.

In sum, the cognitive and interactive model of L2 Chinese reading examined in this dissertation is conceptualized as a model with lower and higher levels of reading processes, with reading acts constantly bi-directional, i.e., moving between these two levels.

1.5 Definition of Key Terms

Terminological usages in the published literature on L2 reading and Chinese reading development are not always consistent. Therefore, for the sake of clarity, definitions of three terms, as used in this dissertation, are given below; and the three terms are character-recognition ability, language background, and L2 reading strategies.

1.5.1 Character-recognition ability. Adapted from definitions of metalinguistic knowledge (Koda, 2005; Kuo & Anderson, 2008; Shu & Anderson, 1999), this study's definition of character-recognition ability is a person's sensitivity to and ability to identify and analyze the orthographic and semantic features of Chinese characters. Originally, metalinguistic knowledge was defined as the ability to identify and analyze the structural features of a language and to use such analyses to facilitate language learning (Koda, 2005; Kuo & Anderson, 2008; Shu & Anderson, 1999). Here, the change in terminology from metalinguistic knowledge to character-recognition ability has two reasons behind it. First, this dissertation concerns a narrower set of metalinguistic knowledge: readers' ability to identify and analyze systematic features in Chinese

characters; and second, it focuses on particular reading processes and their interactions with each other, and places only limited emphasis on readers' ability to apply the accumulated knowledge to language-learning contexts.

1.5.2 Language background. In the current study, language background is specifically limited to a learner's status as either a heritage language learner or a foreign language learner, and such status is treated as a source of background knowledge. Valdés (2000) defines heritage language learners as students who have access to their target language at home and are, to some extent, bilingual in the home language and the mainstream language. Foreign language learners, on the other hand, are those who do not have any familial connection to the target language, and who therefore start from scratch when starting to learn the target language. However, Kondo-Brown (2005) found that the learning advantage conferred by familial connection to the target language disappeared when learners were more than third-generation immigrants, meaning that only native-speaking parents, and not grandparents, were a positive factor in a person developing the target language. Therefore, in the present study, language background is defined according to whether or not the participants' parents are identified as native speakers of Chinese.

1.5.3 L2 Reading strategies. Following Canale and Swain's (1980) framework of communicative competence, Bachman and Palmer (1996) defined strategic competence as a set of metacognitive components responsible for setting reading *goals*, *assessing* the quality of reading comprehension, and *planning* necessary steps to achieve adequate comprehension. Purpura (1997) and Phakiti (2008) subsequently found that reading strategies can be further divided into cognitive and metacognitive types, and that only cognitive strategies directly regulate cognitive reading processing, while metacognitive strategies exert executive functions over cognitive ones. According to Phakiti (2008), L2 reading strategies are conscious, deliberate,

goal-directed mental processes aimed at reaching adequate comprehension during a specific L2 reading task. It should be noted that the definition of L2 reading strategies used in this dissertation differs from the definition of strategic competence discussed in Canale and Swain (1980), in that the present work regards strategies as being activated whenever the readers deem it necessary, and regardless of whether any breakdown in communication and/or comprehension has occurred. When Canale and Swain (1980) first discussed strategic competence, they emphasized its problem-solving nature, meaning that when interlocutors notice a need to repair their utterances, they employ various strategies to compensate for their lack of linguistic and sociolinguistic competence. In other words, L2 strategies are negatively correlated with linguistic and sociolinguistic competence. However, Purpura (1997) and Phakiti (2008) found that different reading strategies correlated with reading comprehension differently. For example, strategies for linking prior knowledge and re-reading were both negatively correlated with reading comprehension, whereas strategies for evaluating comprehension quality and vocabulary knowledge were positively correlated with it (Purpura, 1997). Therefore, following the definitions developed by reading scholars (Afflerbach, Pearson, & Paris, 2008; Phakiti, 2008; Purpura, 1996), this dissertation defines L2 reading strategies as *conscious* and *deliberate* mental processes employed to *facilitate* and *repair* particular reading processes, as determined by *reading purposes*.

1.6 Importance of the Study

The primary goal of this study is to examine the relationships between lower-level processes, higher-level processes, and reading-test performance within a single cognitive model of L2 Chinese reading. In doing so, it operationalizes L2 Chinese reading as being constructed jointly by character-recognition skills, reading comprehension skills, and readers' conscious

attention. Given such a model, reading-test performance is expected to reflect readers' varying degrees of Chinese-character recognition skills, their background knowledge of texts and the target culture, and their reading-strategy use. Although much previous research has touched upon one or more of these factors in L2 Chinese reading, and has recognized that interaction between these factors occurs, reading comprehension is often treated as monolithic, and the complicated cognitive-processing skills involved in forming coherent comprehension and interpretation are therefore overlooked. Moreover, the relationship between Chinese-character recognition and Chinese reading comprehension, both as an L1 and L2, has often been investigated through correlation or multiple regression, implying that their relationship is linear, rather than non-linear and recursive. To address these limitations of the previous research, this study has been designed to investigate subskills involved in character recognition, comprehension, and strategy use, using structural equation modeling (SEM). The resulting model will thus enable a richer description of how reading processes interact with each other, as well as language background's relationship to these reading processes in the domain of L2 Chinese reading.

1.7 Organization of the Dissertation

This dissertation is organized as follows. Chapter 2 reviews the relevant linguistic and cognitive literature on the nature of Chinese orthography, Chinese character-character processing skills, strategy use, and the previously observed differences between Chinese heritage language learners (HLLs) and foreign language learners (FLLs) in the contexts of L1 and L2 Chinese reading. It also presents the dissertation's research questions, and an initial model of L2 Chinese reading. Chapter 3 provides detailed descriptions of the present study's participants, measurement materials, analysis methods, and procedures. Chapter 4 reports the results of the analyses, organized according to the research questions; synthesizes the study's findings in light

of previous research; and suggests possible explanations for trends seen in the data. Finally, Chapter 5 summarizes the findings of the study and its limitations, and suggests avenues for future research.

CHAPTER 2

LITERATURE REVIEW

This chapter focuses primarily on prior studies of L2 Chinese reading, although some literature on L1 Chinese reading has also been included due to its substantial influence on L2 Chinese reading research. The chapter begins with a general introduction to Chinese orthography; this is followed by a review of L2 Chinese reading processes, which include lower-level Chinese character-recognition skills and higher-level strategy use. Although language background is regarded as a higher-level processing variable, character-recognition skills and strategy use can only co-occur with one type of language background, meaning that character-recognition skills and strategy use are nested in the variable of language background. Due to its nesting nature, the variable of language background is discussed in its own section that follows the review of the literature on character-recognition skills and strategy use. After this, the chapter summarizes prior research findings on the reading of Chinese as an L1, as a foreign language (FL), and as a heritage language (HL), with detailed discussion of the component skills involved; and lastly, it sets forth the present study's hypothesized reading-processing model for L2 Chinese reading and the related research questions.

2.1 Chinese Script

The tendency among scholars of L1 and L2 Chinese reading to focus on lower-level processing, especially Chinese-character recognition (e.g., McBride-Chang & Ho, 2000; Perfetti & Dunlap, 2008; Shu & Anderson, 1999; Wang, Spencer, & Xing, 2009), could be related to the weak grapheme-phoneme correspondence (GPC) of Chinese characters, which applies to both the Simplified Chinese characters used in Mainland China and the Traditional Chinese characters used in Taiwan. Unlike Western letters, most Chinese characters encode both phonetic and

semantic information (Ho, Yau, & Au, 2003), with those that include both often being termed compound characters. Sub-character units of these compounds typically encode *phonetic components* on the right-hand side of a character, while semantic information is denoted by a *semantic radical* that is generally located on the left-hand side.

Although beginning Chinese readers are, in addition to Chinese characters, supplied with pinyin – an English-alphabet system indicating the pronunciation of characters – more proficient readers mainly rely on phonetic components when decoding unfamiliar Chinese characters. Pinyin of a character's pronunciation consists of three parts: an initial, a rime, and a tone mark. An initial is a consonant, whereas the basic phonological element of a rime is a vowel, sometimes accompanied by another vowel with additional one or two consonants added before and/or after it (Ye, 2008). Take the pinyin of the character 松 (*sōng* in pinyin, 'pine') as an example. *S* denotes the *initial* consonant of the pronunciation, *ong* denotes the rime, and a tone mark is on the main vowel of the rime, which is *o* in this example.

The challenge of accessing character pronunciation through phonetic components is that such components often only reveal approximate pronunciations of the character, such as its rime. Semantic radicals, on the other hand, offer a relatively more reliable indication of a character's semantic category. For example, the compound character 松 (*sōng* in pinyin, 'pine') is written with the character 木 (*mù* 'trees, woods') as the semantic radical on the left side, and the character 公 (*gōng* 'fair, equitable, public') as the phonetic component on the right side, providing an approximate pronunciation of the character 松 (*sōng*) – in this case, the rime *ōng*.

There are approximately 800 phonetic components and 200 semantic radicals (Ho, Yau, & Au, 2003). As suggested by the example given in the previous paragraph, many of the phonetic components and the semantic radicals can also be stand-alone characters. On the other

hand, some of the phonetic and semantic units are not stand-alone characters, which are termed *stroke patterns*, and the stroke patterns, as well as stand-alone characters, are assigned to particular positions within characters that readers are required to notice. The combining of phonetic components and semantic radicals into compound characters commonly follows one of two patterns: left-right and top-bottom (Ho, Yau, & Au, 2003; Shen & Bear, 2000). For instance, the regular positions of semantic radicals 木, 欠, 艹, and 灬 are on the left, the right, the top, and the bottom of characters, respectively. Sub-character units with correct forms and positional regularity are fundamental to be recognized as correct Chinese characters. In the example 杉 (*shān*, 'Cunninghamia, China Fir, a conifer'), the semantic radical 木 and the phonetic component 彡 occupy their regular positions. To form another character that combines the same phonetic component with the semantic radical 衤(denoting garment, clothes and clothing), the correct configuration is 衫 (*shān*, 'garment, jacket with open slits in place of sleeves'), again with the semantic radical on the left and the phonetic component on the right. To place 彡 on the left, the top, or the bottom violates the rules of positional regularity, even though the form of the stroke pattern is correct.

Because a Chinese character is often a morpheme, a Chinese word can consist of one to six Chinese characters, but most Chinese words comprise two (Shen & Bear, 2000). A given morpheme often has more than one meaning, but when it is combined with one or more other morphemes, the resultant word has only one meaning. Take 松树 (*sōngshù*, 'pine tree') as an example. The word is composed of two morphemes, 松 (*sōng*, 'pine') and 树 (*shù*, 'trees'). Similarly, 松木 (*sōngmù*), which consists of the morphemes 'pine tree' and 'trees, wood', means 'pine wood'; and the words 杉树 (*shānshù*) and 衬衫 (*chènshān*) mean 'China Fir tree' and

'shirt, or blouse', respectively. In these three words, 松树, 松木, and 杉树, the semantic radical of all six characters is 木 ('trees, woods'), indicating the broad semantic category, but the more refined semantic information is provided by characters, or morphemes, after readers have recognized that 松树 and 杉树 refer in some way to trees, due to the presence of the morpheme 树. Similarly, 衬 ('[of garments] against the skin, to line, lining') and 衫 ('garment, jacket with open slits in place of sleeves') provide more semantic information than the semantic radical 衤 for clothes and clothing does, when one is seeking to decode the meaning of the word 衬衫 ('shirt, blouse').

2.2 L2 Chinese Reading Processes

In light of the aforementioned descriptions of Chinese characters and writing system, I can now introduce the component skills involved in reading Chinese texts. As noted earlier, many scholars regard reading as an interactive process that involves a combination of lower-level and higher-level processes, multiple reading skills, textual information, and a reader's prior knowledge (for reviews, see Grabe, 2009; Hudson, 2007, Stanovich, 2000). *Lower-level processing* denotes word recognition skills, including phonology processing, orthography processing, morphology processing, and lexical access, whereas *higher-level processing* means comprehension skills, such as building text and situation models of texts, skimming, scanning for main ideas, making inferences, and predicting incoming text content (Grabe, 2009; Hudson, 2007).

Regardless of whether researchers are examining the reading of Chinese as an L1, L2, or HL, they have strongly emphasized lower-level processing (e.g., for bilingual L1 children: McBride-Chang & Kail, 2002; Shu, et al., 2003; Tong & McBride-Chang, 2010; for Chinese HLLs and FLLs: Everson, 1998; Lü & Koda, 2011; Shen & Ke, 2007). Lower-level processing is

a necessary aspect of reading comprehension; however, it is insufficient in owe of itself. In other words, reading comprehension can be driven by word recognition, but automatic comprehension is not guaranteed by effective lower-level processes alone. Moreover, a person's mental resources can be consciously relocated to monitor and redress interruptions to her/his comprehension (Hudson, 2007), and such conscious effort is often defined as reading-strategy use (Grabe, 2009; Hudson, 2007; Phakiti, 2003, 2008; Purpura, 1997). In sum, this dissertation aims to arrive at a more interactive model of L2 Chinese reading, by including both lower- and higher-level processing and taking into consideration the interactions between processing skills, texts, and readers.

2.2.1 Lower-level processing. Word-recognition skills are regarded as lower-level because they are usually the first challenges faced by beginning readers; and it is noteworthy that the label "lower" does not mean that these skills are easier to master than higher-level ones. When encountering any written script, readers must be able to make connections between the written symbols (graphemes), their oral pronunciation (phonemes), and words' meanings. The faster and more automatically such linkages are made, the lower the processing load on the reader's working memory, thus releasing the remaining working memory for higher-level processing (Grabe, 2009; Hudson, 2007). Investigations of Chinese word recognition have tended to focus on three main categories of lower-level processing skills: orthography, phonology, and morphology (Perfetti & Tan, 1999). However, this dissertation considers the lower-level skills particularly related to reading Chinese characters, which are ortho-phonology, ortho-semantics, and morphology. Ortho-phonology and ortho-semantic skills are included to better reflect the decoding process of Chinese-character recognition and its relationship with L2 Chinese reading. Investigating both of these skills allows me to examine how test-takers access

sub-character levels of information for character recognition. Morphology processing, meanwhile, provides more refined and accurate information about a given word than semantic radical does (Tong & McBride-Chang, 2010).

2.2.2 Higher-level processing. In contrast to lower-level approaches' usual focus on the crucial role of word-recognition skills in reading comprehension, higher-level ones emphasize the cognitive activities involved in reading comprehension, such as integrating printed information and weaving pieces of this information into a coherent understanding of a passage (Grabe, 2009; Hudson, 2007). In addition to forming adequate comprehension, strategy use is another important factor for fluent reading (Grabe, 2009; Grabe & Jiang, 2014; Hudson, 2007; Phakiti, 2003, 2008; Purpura, 1997; Schoonen, Hulstijn, & Bossers, 1998).When individuals' comprehension levels are sufficient to the reading tasks at hand, their reading-processing is so rapid and automatic that they are aware neither of it nor of the cognitive activities involved – collectively, *reading skills* – which include selecting information, analyzing, reconstructing, and evaluating newly formed understandings. On the other hand, comprehension that is insufficient to the reading purpose necessitates conscious repair, and the process is no longer automatic, but controlled (Cohen, 2014; O'Malley & Chamot, 1990; Oxford, 2011). Two higher-level variables, strategy use and language background, are discussed in the following two sections.

2.2.2.1 Strategy use in L2 reading. Prior to discussing the possible role of strategy use in L2 Chinese reading, I will first introduce relevant L2 studies that have investigated the importance of strategy use as a higher-level process; then, prior studies that have focused on Chinese-reading strategies using other theoretical frameworks, such as Bernhardt's L2 reading processes (1991) in the studies of Everson & Ke (1997) and Lee-Thompson (2008) and

Oxford's L2 learning strategies (1990) in Sung's study (2011) will be dealt with in the following three sections of Chinese reading as an L1, L2, and HL.

Based on information-processing theories, when the automatic process described above is interrupted, cognitive activities are activated in an attempt to best utilize the limited-capacity working memory to process the information most crucial to the reading purpose. The main distinction between reading-processing skills and reading strategies therefore relates to the concepts of automaticity and consciousness. In other words, when cognitive activities are employed in automatic reading processing, with little conscious effort required to facilitate reading comprehension, these activities are referred to as reading skills; but when the same mental activities are activated in an attempt to mend reading comprehension, they are defined as reading strategies (Cohen, 2014; Grabe, 2009; Phakiti, 2003; Purpura, 1997; Urquhart & Weir, 1998).

When investigating the relationship between reading-test performance and strategy use, Purpura (1997, 1999) and Phakiti (2003, 2008) also examined the construct of strategy use while taking reading tests. Taken as a whole, their findings suggest that reading strategies can be divided into two categories: cognitive and metacognitive. Cognitive strategies are mental activities used for *comprehending* passages (e.g., identifying main ideas and authors' attitudes, translation, prediction, and inferencing), *memorizing* information without generating fresh understanding of the targeted part (such as by rereading, repeating, note taking, or paraphrasing), and *retrieving* relevant prior knowledge (such as one's own prior experience, or grammatical rules). Metacognitive strategies, in contrast, are mental activities used for *planning* (such as goal-setting and arriving at overviews of tasks), *monitoring* (e.g., noticing comprehension failures or

errors, double-checking comprehension), and *evaluating* comprehension (by assessing levels of difficulty, self-testing, and evaluating test performance and accuracy).

Adopting Purpura's (1999) construct of cognitive and metacognitive strategies, Phakiti (2008) investigated the relationship between strategic competence and reading-test performance. He created a reading test divided into two parts, a gap-filling test targeting lexical, structural, and discourse knowledge, and a multiple-choice reading-comprehension test targeting the identification of main ideas, details and inferences. The results suggested a stable construct of strategy-use for test-takers with differing test performances and across two test settings (i.e., mid-term vs. final exams), although the factor loadings of these strategies varied for examinees with higher and lower test scores, and for the two test settings. Both studies (Phakiti, 2008; Purpura, 1999) support a unidirectional reading model in which metacognitive strategies regulate cognitive-strategy use to facilitate performance on the fill-in-blank test, which in turn contributes to performance on the passage-comprehension test.

In terms of the relationship between strategy use and reading-test performance, the former has been found to explain between 12% and 30% of the total variance in the latter (i.e., 12% for cognitive strategies and 22% for metacognitive strategies in Phakiti, 2003; 30% of cognitive strategies in Test 1 in Phakiti, 2008; and 12.5% for memory, retrieval, and monitoring strategies together in Song, 2005). Although a positive correlation between strategy use and reading-test performance was also found (Phakiti, 2003, 2008; Purpura, 1999; Song, 2005), the relationships between individual strategies and test performance were more profound. Purpura (1999), for instance, identified a positive relationship between the use of metacognitive strategies and reading-test scores, as well as a negative relationship between the use of memory strategies and scores. Based on data collected from 161 participants using Purpura's (1999) strategy

questionnaire, Song (2005) found similar complicated relationships between strategies and reading-test performance. Her exploratory factor analysis and regression analysis results indicated that monitoring and retrieval strategies were positively correlated with reading-test performance, while memory strategies were again negatively correlated with it, confirming Purpura's original finding.

2.2.3 Language background. In this study, language background serves as a grouping variable, differentiating readers into two broad categories – FLLs and HLLs – along with some sub-categories of Chinese HLLs. It is important to differentiate HLLs from FLLs in foreign-language classrooms in the United States for three reasons. First, many HLLs have early and ample exposure to aural input in the target language in natural settings, such as in their homes and within communities where the language is widely spoken, resulting in better pronunciation, larger vocabulary sizes, and better informal registers than FLLs have (Kagan & Dillon, 2008; Lynch, 2003; Montrul, 2010, 2013; O'Grady, Lee, & Lee, 2011). Second, HLLs are likely to exhibit better morphosyntactic intuition, although it has been noted that their morphosyntactic knowledge is more procedural and less declarative, due to the natural contexts in which they acquire the target language (Montrul, 2012; Zhang, 2014). Thus, HLLs may not be able to transfer their procedural morphosyntactic understanding into test performance, to the extent that testing requires declarative knowledge of the target language that has been developed via formal instruction in it (Montrul, 2012). Third, HLLs' dominant languages begin to shift to the mainstream language at school age and increasingly so in subsequent years, leading to incomplete acquisition of their home language (Montrul, 2010; O'Grady, Lee, & Lee, 2011); indeed, many HLLs do not develop age-appropriate language proficiency in terms of vocabulary, grammar, and literacy in their home languages, and only a handful are able to reach advanced

proficiency in all four language skills, which are listening, speaking, reading, and writing (Benmamoun, Montrul, & Polinsky, 2010; Montrul, 2010, 2012, 2013). Accordingly, even though HLLs exhibit some linguistic advantages over FLLs, HLLs' utterances often incorporate vocabulary and grammar that would be more expected in young children.

In addition to noting their differences from FLLs, it is also important to recognize that HLLs are a heterogeneous and diverse group (Valdés, 2005). Kondo-Brown (2005) investigated the language profiles of four groups of Japanese-language learners, three consisting of HLLs and one of FLLs, and found that the only learners who differed significantly from the others were HLLs with more than one Japanese-speaking parent. Such a result indicates that the importance of parents' language use at home may be higher than that of other caregivers, such as grandparents and other relatives.

In addition to language use at home, diversity among HLLs can be identified in terms of their target languages' various regional vernaculars. Usually, the concept of "dialect" is socially constructed, rather than linguistically defined (Ortega, 2013; Valdés, 2005). Noticing the diversity brought to language classes by HLLs with various dialect backgrounds, researchers have found important differences in language use between the dominant dialects being taught and other less-dominant dialects spoken in these students' homes (for Arabic, see Albirini, Benmamoun, & Saadah, 2011; for Chinese, Wiley, 2008 and Wong & Xiao, 2010; and for Spanish, Villa, 1996).

For her research on Chinese HLLs in the United States, He followed Valdés (2000) in defining a Chinese HLL as "a language student who is raised in a home where Chinese is spoken and who speaks or at least understands the language and is to some degree bilingual in Chinese and in English," and is "English-dominant with no or limited reading/writing ability in Chinese"

(He, 2006, p. 1). On the one hand, these definitions create a space for Chinese HLLs from less-dominant dialects in language classes, but on the other, they ignore the vast linguistic differences between Chinese dialects.

The overlooking of differences among such dialects is related to the fact that the term "Chinese," as applied to language, is an overarching concept including all the dialects spoken by ethnic Han people in China (DeFrancis, 1984). However, the English term "dialect" is misleading in this case, insofar as it is a loose translation of *fāngyán* in Mandarin, which in fact means regional speech (DeFrancis, 1984); and the difference between one region's *fāngyán* and another's can be tremendous. Chao (1947, p. 5) describes the difference between Mandarin and Cantonese as similar to that between English and Dutch, despite Cantonese being regarded as a Chinese dialect rather than as a separate language. Most often, the "Chinese" taught in American Chinese-language programs is Mandarin, the dominant dialect, but many other Chinese dialects are actually not intelligible to Mandarin speakers. Therefore, Chinese HLLs whose home *fāngyán* are not Mandarin tend to have difficulties in comprehending spoken Mandarin in class.

The issue of low mutual intelligibility between *fāngyán* is further complicated by the existence of two written scripts used by Chinese people, the Simplified characters and the Traditional characters. For example, a Cantonese speaker might fail to comprehend Mandarin directions given orally, and yet be able to read the class textbook.

To supplement prior definitions and reflect the low mutual intelligibility between *fāngyán*, He (2006, p. 3) further divided Chinese HLLs into two categories with three subgroups in each, as follows:[1]

1. Mandarin is the learner's home *fāngyán* or is comprehensible to home *fāngyán* speakers

[1] The term "dialect" in the original description has been replaced with *fāngyán* to reflect the low level of mutual intelligibility between Mandarin and the other regional speeches.

   a. The classroom script is the same as the home script (i.e., both use Traditional or Simplified characters)
   b. The classroom script differs from the home script (i.e., Traditional is used in class and Simplified at home, or vice versa)
   c. There is no home literacy in Chinese
2. Mandarin is unintelligible to home *fāngyán* speakers
   a. The classroom script is the same as the home script (i.e., both use Traditional or Simplified characters)
   b. The classroom script differs from home script (i.e., Traditional is used in class and Simplified at home, or vice versa)
   c. There is no home literacy in Chinese

In comparison to HLLs of other languages, the typical HLLs of Chinese are those whose home *fāngyán* is Mandarin and whose home script is the Simplified character script would have the greatest advantage because their spoken vernacular and character script are the same as the ones taught and valued in most Chinese language programs. On the other hand, when being classified as Chinese HLLs, other *fāngyán* speakers often experience difficulties related to their limited Mandarin exposure (Kelleher, 2008; Li & Duff, 2008; Wiley, 2008), and researchers have suggested that other *fāngyán*-speaking Chinese learners should be recognized as a subgroup of Chinese HLLs who exhibit different learning characteristics than their Mandarin-speaking counterparts (Man-HLLs) (Lee, 1996; Li & Duff, 2008; McGinnis, 1996; Wiley, 2008; Wong & Xiao, 2010). Scholars (Hendryx, 2008; Li & Duff, 2008) have suggested that other *fāngyán*-speaking Chinese HLLs are likely to have lower oral skills and better literacy skills than Man-HLLs do.

Among other *fāngyán*-speaking Chinese HLLs, Cantonese is one of the most commonly spoken *fāngyán* in countries in North America, the Pacific, and Southeast Asia (Bradley, 1992; Norman, 1988), and was the first *fāngyán* formally taught in the United States (He, 2008). Most Cantonese-speaking Chinese HLLs (Can-HLLs) are descendants of people from Guangdong Province and the surrounding areas such as Hong Kong and Macau, which have a long history of migration to other countries. Considering the differences within the group of Chinese HLLs and the large population of Can-HLLs enrolled in Chinese language programs in North America, this study considers Man-HLLs to be the most advantaged subgroup and chooses Can-HLLs as a *fāngyán*-speaking subgroup to represent Chinese HLLs' intragroup variation.

In light of the aforementioned characteristics of Chinese HLLs, it was decided that the present study should examine language background as a higher-level processing variable. Specifically, this would be achieved through including three groups of Chinese-language learners, 1) a group of Man-HLLs whose parents are all Mandarin native speakers, 2) a group of Can-HLLs whose parents are all Cantonese native speakers, and 3) a group of English-speaking FLLs whose parents are all English native speakers. It is expected that, among all types of Chinese-language learners, these Man-HLLs would have the most access to background knowledge of the target language, both linguistically and culturally, while the FLLs would likely have the least access.

2.3 Reading Chinese as a First Language

Many scholars who investigate Chinese reading development often assume that advanced character-recognition processing skills initiate rapid and automatic character recognition, leading to successful Chinese-character reading, and thence to sentence- and passage-reading comprehension (e.g., Cheung, Chan, & Chong, 2007; Koda, Lü, & Zhang, 2008). A parallel line

of investigations of both L1 and L2 Chinese reading, likewise focused on lower-level approaches, uses character-recognition processing skills as indicators of Chinese reading ability (e.g., Cheung et al., 2007; Koda, 2002; Tong & McBride-Chang, 2010).

Because the focus of the present research is on readers' abilities to connect phonemes, graphemes, and meanings when processing Chinese texts, only orthography and morphology processing skills are included in its investigation of lower-level reading processing. It also further divides orthography processing skills into ortho-phonology and ortho-semantics skills, to reflect the characteristics of Chinese orthography.

Several scholars have found orthography and morphology processing skills to be significantly correlated with L1 Chinese character reading comprehension (Cheung et al., 2007; Li, Anderson, Nagy, & Zhang, 2002; Shu, et al., 2003; Tong, 2008). In Li et al.'s (2002) investigation of Chinese first and fourth graders' reading development, the combination of phonology-, morpheme-, and ortho-semantics processing skills explained approximately 40% of the variance in L1 comprehension (39.3% for first grade and 46.9% for fourth grade).[2] The same research also suggest that morphology processing skills play a more important role than phonology knowledge does in reading comprehension, and that as L1 readers grow older, morphology processing skills become more important for reading comprehension (explaining 23% of comprehension variance for first graders, but 38% for fourth graders).

Although Li et al. (2002) did examine morpheme- and ortho-semantics processing abilities, it should be noted that they grouped both of these skill sets into a single category, "morphology processing skills", and therefore did not take into account the difference between the semantic information provided by a Chinese character and a semantic radical. Especially in

[2] The number of first-grade participants who completed each instrument ranged from 84 to 303, and the number of fourth-grade participants from 109 to 305.

light of the fact that studies investigating the role of character-recognition skills in Chinese reading development often have divergent views regarding morphology processing skills, it seems advisable to consider radical knowledge separately from such skills, even though both provide semantic information in the decoding process.

In a more refined study of character-recognition development, Tong (2008) set 12 tasks to explore the evolving constructs of phonology, orthography, morphology, and sub-character processing skills.[3] The participants were recruited from three age groups: kindergarteners (*n* = 199) as pre-readers, second graders (*n* = 172) as beginning readers, and fifth graders (*n* = 165) as advanced readers. The results of SEM suggested a strong influence of print-experience on the development of Chinese-character recognition. For pre-readers, the four variables clustered into two groups, one tested orally and the other one tested through print. For beginning readers, a four-construct model emerged, indicating their burgeoning abilities to process the information from morphology, orthography, and sub-character configuration to achieve Chinese-character recognition. For advanced readers, a two-construct model was supported, in which morphology, orthography, and sub-character skills functioned similarly, and these three types of skills interacted with each other to a large extent, and only phonology skills functioned differently from the other three types. Following from my general discussions of lower- and higher-level processes, this section and the next comprise a more detailed examination of these processes with respect to reading Chinese as an L1, an FL, and an HL; and in each of these sections, character-recognition skills are discussed first, followed by strategy use.

[3] In Tong's (2008) study, phonology skills were measured using a syllable- and syllable onset deletion task, a rhyme production task, and a tone detection/discrimination task. Morphology skills were measured via a homographic discrimination task, a morphological-construction and homophone-production task, and a morpho-grammar derivation task. Orthography-processing skills were measured using a stroke-pattern position task, a visual configuration judgment task, and a character decision task. Sub-character skills were measured by an ortho-semantic awareness task, an ortho-phonetic awareness task, and a semantic-picture-mapping task.

2.3.1 Ortho-phonology processing skills. Chinese L1 children become aware of the orthographic composition and the functions of phonetic components and radicals as early as the age of six (Li et al., 2002; Shu & Anderson, 1999; Tan & Perfetti, 1997; Tong, 2008). Moreover, they are able to extract phonological information before being explicitly taught about the underlying rules (Cheung & Ng, 2003; Tzeng, 2002). Shu and Anderson (1999) observed Chinese children's ability to analyze phonetic components when encountering unfamiliar characters, indicating the early importance of ortho-phonology processing skills.

In a more recent regression-analysis study, Cheung et al. (2007) found ortho-phonology processing skills to be significant predictors of reading comprehension and the ability to read aloud among 88 Chinese L1 children aged from 9 to 11.6 years old. After controlling for the effects of age, nonverbal IQ, and working memory, these children's ortho-phonology processing skills contributed approximately 5% of the unique variance in Chinese reading comprehension and 7% of the unique variance in Chinese reading aloud. In the same study, ortho-phonology processing skills also shared 10% of total variance with Chinese reading comprehension, based on simple correlation.

Studies of bilingual Chinese children have revealed that, while the contribution of ortho-phonology processing skills to Chinese reading comprehension is significant, it is not as strong as its contribution to English reading comprehension (Tong, 2008; Tong & McBride-Chang, 2010). Tong and McBride-Chang (2010) suggested that this difference was related to the lower GPC of Chinese, which might lead the processing of Chinese characters to depend less on phonetic components than the processing English words does.

2.3.2 Ortho-semantics processing skills. Research suggests that Chinese children actively apply their radical knowledge when decoding unfamiliar characters to determine their

meaning, and when reading sentences (Cheung et al., 2007; Feldman & Siok, 1999; Li et al., 2002; Nagy, Kuo-Kealoha, Li, Anderson, & Chen, 2002; Shu & Anderson, 1999). When asked to choose the character most similar to a target character, the Chinese children studied by Chen (2003) tended to select one with the same radical as the target character, rather than the same phonological component. This suggests that Chinese children prefer ortho-semantic cues over ortho-phonological ones in character recognition. Other studies have confirmed the important roles of semantic radicals in character recognition and Chinese reading comprehension (Cheung et al., 2007; Li et al, 2002). Cheung and colleagues' (2007) abovementioned investigation of the contributions of ortho-phonological and ortho-semantics processing skills to Chinese reading aloud and Chinese reading comprehension found that the latter type of skills explained 8% and 7% of their unique variance, respectively, after controlling for the variables of age, nonverbal IQ, and working memory. In addition, ortho-semantics processing skills shared 12% of total variance with Chinese reading comprehension, based on simple correlation. Although both ortho-phonological and ortho-semantics processing skills seemed to be of limited utility for predicting Chinese reading comprehension, their contributions (based on the results of multiple regression, controlled for age) were only less than the contribution of nonverbal IQ, and larger than the contribution of working memory.

2.3.3 Morphology processing skills. Morphology processing skills, which are the skills used to determine the meaning of unfamiliar characters and words in contexts, have been described as strongly predictive of Chinese reading comprehension (Ku, 2000; Kuo & Anderson, 2008; Li et al., 2002; Tong, 2008). Ku (2000) reported that for L1 Chinese second to sixth graders, morphology processing skills[4] were significantly correlated with both vocabulary

[4] While morphology skills were examined orally in Tong's 2008 study, paper-and-pencil testing was used by Ku (2000). Therefore, extra caution is necessary when comparing morphology skills data across the two studies.

knowledge (r_{xy} between 0.5 and 0.6, $p < 0.05$) and Chinese reading comprehension (r_{xy} between 0.6 and 0.75, $p < 0.05$), and that these skills continued to change across the entire age range of the children in the study.

However, based on the results of hierarchical regression analyses in the same study, Ku (2000) also found that the unique contribution of morphology processing skills to Chinese reading comprehension decreases markedly over time: from 32.1% in second grade to 6.5% in sixth grade, after controlling for vocabulary knowledge. Given the improvement in the participants' performance of morphology tasks across grades, the decreasing unique variance of morphology skills may be explained by a large covariance with some other general comprehension ability, such as vocabulary knowledge or the comprehension ability required for the task of re-ordering sentences. Ku's explanation based on covariance with other skills was supported by Tong's aforementioned 2008 study, in which morphology skills were grouped with or correlated to other lower-level processing skills.

Similar covariance between morphology skills and other reading skills were found by Hu's (2013) study of third and fifth graders, which found that oral morphology skills were significantly correlated with vocabulary knowledge, phonology skills, and Chinese character recognition (r_{xy} between 0.3 and 0.45, $p < 0.05$). Such skills also explained approximately 2% and 5% of unique variance in Chinese reading comprehension for the two age groups, respectively.

2.3.4 Strategy use. In an investigation of the relationship between metacognitive strategy use and reading comprehension among 304 Chinese sixth graders, Chan and Law (2003) used a metacognitive-strategy questionnaire covering cognitive strategies, self-regulation, and Chinese comprehension. They observed that those children with better comprehension scored higher on

the other sections of the questionnaire than those with lower comprehension levels; and, using hierarchical multiple regression, they found a positive but insignificant correlation between strategy-use and Chinese reading comprehension. In conclusion, Chan and Law (2003) suggested that strategy-use facilitates literal and inferential comprehension for L1 Chinese children. However, the authors did not analyze cognitive strategies and self-regulation separately, and their complete metacognitive strategy questionnaire was not provided in the article. On the whole, it seems that a more refined categorization would be better able to uncover the construct of strategy use.

2.4 Reading Chinese as a Foreign Language

English-speaking learners of Chinese as a foreign language (CFLLs) often struggle with its unfamiliar orthographic features (Everson, 2011), and when decoding Chinese characters they may rely too heavily on ortho-phonology processing skills developed through English reading (Perfetti & Dunlap, 2008). According to Everson (2011), additional challenges for CFLLs include their limited exposure to Chinese texts include relatively small vocabulary size, limited spoken fluency when beginning to read, and a lack of background knowledge of the target language, target culture, and discourse structure.

To investigate the influences of language background on Chinese-character recognition and production, Ke (1998) recruited 59 English-dominant CFLLs and 84 Chinese HLLs enrolled in first-year Chinese courses at nine American universities. At the end of the first year, the participants completed a character-recognition task and a character-production task. Their scores on the two tasks showed no significant group differences. Ke concluded that the home linguistic input that the Chinese HLLs received was insufficient to allow them to significantly outperform CFLLs in terms of Chinese-character development.

In another study, Everson (1998) recruited 20 CFLLs who had studied Chinese for two complete years and whose native languages used alphabetic orthographies. The participants were asked to pronounce Chinese characters and provide their meanings in English. Much like L1 Chinese children (Tong & McBride-Chang, 2010), these beginning CFLLs were able to connect pronunciation and meaning to the orthographic representations. Everson therefore suggested that beginning CFLLs are able to establish GPC even with low-GPC Chinese characters, and that they do not rely on rote memorization to learn Chinese characters.

When reviewing L2 Chinese learning-strategy research, Jiang and Cohen (2012) found that most of the examined strategies related to identifying character- and word meanings. Shen (2013) divided character-learning strategies into two categories: cognitive and metacognitive. The former are related to *memorizing* and *analyzing* character configuration, and *applying rules* to the learning of new characters. Metacognitive strategies, on the other hand, are related to *evaluating* learning errors and *planning* future learning accordingly. In other words, character-learning strategies involve far more than character-recognition processing skills. Likewise, Li (1998) found – based on qualitative data collected from open-ended questionnaires, structured interviews, and reading/thinking-aloud tasks – that beginning CFLLs took account of orthographic similarity, syntactic acceptability, and semantic acceptability when deciphering unfamiliar Chinese words.

Two other studies based on participants' verbal reports following thinking-aloud protocols (Everson & Ke, 1997; Lee-Thompson, 2008) found that intermediate-level CFLLs were able to apply various cognitive and metacognitive reading strategies, such as making assumptions about a text's genre and possible topics, and making inferences based on a limited number of known words or context clues. Regarding the use of cognitive strategies in particular,

Everson and Ke (1997) found that more advanced participants demonstrated better phonological/grapheme decoding, word recognition, and morphology processing skills, and were able to apply more reading strategies, and apply them more effectively, than less advanced participants were. In terms of metacognitive strategies, the advanced participants were able to monitor their comprehension, identify problems with it, and make adjustments to repair these problems by using strategies such as focusing on the article title or repeated phrases, rereading, and assessing the importance of unfamiliar words. Along the same lines, participants in Lee-Thompson's (2008) study frequently reported purposefully monitoring and repairing their word-level comprehension, implying that word recognition may be at the center of lower- and higher-level processing for intermediate CFLLs.

Some pedagogy studies have explored the advantages of instruction in radical knowledge (e.g., Shen & Ke, 2007; Wong, 2011; Xu, Chang, & Perfetti, 2014) and found that grouping characters with the same radicals (Xu et al., 2014), explicit instruction in analyzing radical meanings (Shen & Ke, 2007; Wong, 2011), and applying radical knowledge to word recognition (Shen & Ke, 2007) all facilitate the learning and retention of character knowledge among both CFLLs and Chinese HLLs.

As previously discussed, CFLLs are challenged by the unfamiliarity of Chinese characters, Chinese's weak GPC, and their own lack of background knowledge in the target language, target culture, and discourse structure (Everson, 2011; Lee-Thompson, 2008). Although they are able to systematically apply character knowledge (Shen & Ke, 2007; Wong, 2011; Xu et al., 2014), they are likely to rely more on ortho-phonological cues than on ortho-semantic ones (Everson & Ke, 1997; Lee-Thompson, 2008; Wong, 2011). Additionally, cognitive and metacognitive strategies are both implemented during Chinese reading processing;

and much like English-language learners, less advanced CFLLs adopt fewer reading strategies, and do so with less success, than advanced CFLLs do (Everson & Ke, 1997).

2.5 Reading Chinese as a Heritage Language

Chinese HLLs in English-speaking countries usually develop English literacy at school while continuing to develop their Mandarin literacy at home and/or at weekend schools (Koda, Zhang, & Yang, 2008; Xiao, 2008). Studies indicate that most Mandarin literacy activities are related to weekend schooling, and the most common activities include completing Chinese homework, reading Chinese textbooks, and preparing for Chinese vocabulary quizzes (Lü & Koda, 2011; Xiao, 2008; Zhang & Koda, 2011). Activities not related to weekend school, such as visiting bookstores or libraries, or reading Mandarin newspapers alone and with parents, were less common. Due to their limited amount of linguistic input and literacy support, Chinese HLLs are likely to encounter more difficulties in developing their lower-level processing in Chinese reading than L1 children. Similar findings were reported in Mu's (2014) study based on survey data collected from 230 Chinese HLLs: socioeconomic status, contacts with Chinese-speaking relatives, age of immigration, and Chinese use at home together explained 74% of the variance in self-reported Mandarin proficiency. However, Mu's inclusion of formal education in Mandarin did not explain additional variance in the multiple-regression formula.

Other studies indicate that, due to limited written input and literacy support and the weak GPC of Chinese orthography, early exposure to Mandarin may not be as supportive a factor in literacy development for young Man-HLLs as for young HLLs of other languages (e.g., Everson, 2011; Ke, 1998; Koda, Lü, & Zhang, 2008; Lü & Koda, 2011; Xiao, 2006). To uncover how young Chinese HLLs develop bi-literacy, Koda, Zhang, and Yang (2008) investigated the correlations between the reading-comprehension skills, on the one hand, and the ortho-semantic

and morphology processing skills of third- to fifth-grade Chinese HLLs (*n* = 23, *n* = 20, *n* = 16, for Grades 3, 4, and 5 respectively). The results did not show that early exposure to Chinese script helped develop Chinese reading skills, either in terms of character-recognition skills or in terms of Chinese reading comprehension. Based on the results of 10 tasks that were assigned to all three age groups, limited improvement in ortho-semantic and morphology processing skills could be found across age groups; comprehension-task performance, however, was significantly and positively correlated with increases in the participants' ages. That being said, the descriptive statistics of the total comprehension scores suggest that this age benefit was relatively small, with accuracy rates of 0.31 (*SD* = 0.19), 0.32 (*SD* = 0.18), and 0.48 (*SD* = 0.27) for the third-, fourth-, and fifth-grade groups respectively. Taken as a whole, Koda, Zhang, and Yang's results suggest that early exposure to Chinese texts may aid certain morphology skills for Chinese HLLs, but that this benefit is not significant. The study findings indicate interrupted lower-level processing, which would further suggest that English-dominant Chinese HLLs need to rely heavily on other higher-level skills and reading strategies which are developed through English reading to repair their Chinese comprehension, just as CFLLs do (Koda, Zhang, & Yang, 2008).

Similar findings were reported by Lü and Koda (2011), who investigated how home literacy support facilitates Chinese-character recognition among young Chinese heritage children. The authors measured their 37 participants' Chinese-character recognition abilities using oral-vocabulary, ortho-phonological, and character-naming tasks. They then divided the participants into two groups: those who read Chinese texts with their parents or by themselves more than twice a week were defined as High-in-Chinese (*n* = 10), and those who did so less than twice a week were defined as Low-in-Chinese (*n* = 27). Although the High group performed

better at Chinese oral vocabulary knowledge and real Chinese character naming than the Low group did, the difference was not statistically significant.

These two studies (Koda, Zhang, & Yang, 2008; Lü & Koda, 2011) suggest some features of reading Chinese as an HL. First, it is clear that individuals' development of age-appropriate Chinese reading proficiency requires more intense and consistent home literacy support than doing Chinese homework twice a week can provide. Second, although there was a positive relationship between character naming and home literacy support, the relationships among character-recognition skills, age, and home literacy support were less clear, and may be insignificant. Third, in order to compensate for their interrupted lower-level processing in reading Chinese, English-dominant Chinese HLLs may apply higher-level skills and strategies that have been developed through English reading.

To better understand the influence of language background on strategy use when learning Chinese characters, Sung (2011) conducted a multivariate analysis of variance (MANOVA) study on 61 Chinese HLLs and 73 CFLLs using Oxford's (1990) Strategy Inventory for Language Learning (SILL). The results suggested that those who had learned more than two FLs made greater use of cognitive strategies than did those who had learned fewer than two FLs[5] ($p < 0.001$, $\eta^2 = 0.12$), and that CFLLs made more use of metacognitive strategies than Chinese HLLs did ($p < 0.001$, $\eta^2 = 0.11$). Sung concluded that Chinese HLLs were better at utilizing their familial resources through social strategies in their HL learning, but also less likely to adopt cognitive and metacognitive strategies, precisely because their main purpose in learning Chinese was to strengthen their familial and cultural connections, and not to promote their literacy skills.

[5] Four participants who had learned Chinese as their L2 were excluded from the analyses in the study.

2.6 Proposed Model of Reading Chinese

The present dissertation conceives of L2 Chinese users as active test-takers who have developed various reading skills to facilitate their test performance, rather than as passive processors who are bounded by their limited vocabulary size. In other words, whenever automatic character recognition does *not* occur, these readers are able to consciously and actively retrieve relevant prior knowledge and regulate their limited mental resources in ways that facilitate reading comprehension. Accordingly, this investigation focuses on the contribution of three kinds of reading-processes to reading-test performance. These are 1) character-recognition ability, to tap into lower-level processing, and 2) reading strategies and 3) reading comprehension ability, representing two distinct kinds of higher-level processing. In addition to the three kinds of reading skills, language background is operationalized as a source of background knowledge; and it serves as a grouping variable when assessing its effect on L2 Chinese reading.

Three common sub-tests have been chosen to assess L2 Chinese reading comprehension ability: a multiple-choice grammar test, a fill-in-the-blank cloze test, and a multiple-choice passage-comprehension test. Guided by prior literature regarding Chinese-character development among monolingual and bilingual children (Cheung et al., 2007; Li et al., 2002, Tong, 2008), this dissertation investigates how ortho-phonology, ortho-semantics, and morphology processing skills facilitate character recognition, and how it later contributes to reading comprehension ability.

Also as suggested by previous research (Purpura, 1999; Phakiti, 2008), strategy use is held to be composed of metacognitive and cognitive strategies. The former regulate the latter,

which in turn affect readers' reading comprehension, and thus their general reading-test performance.

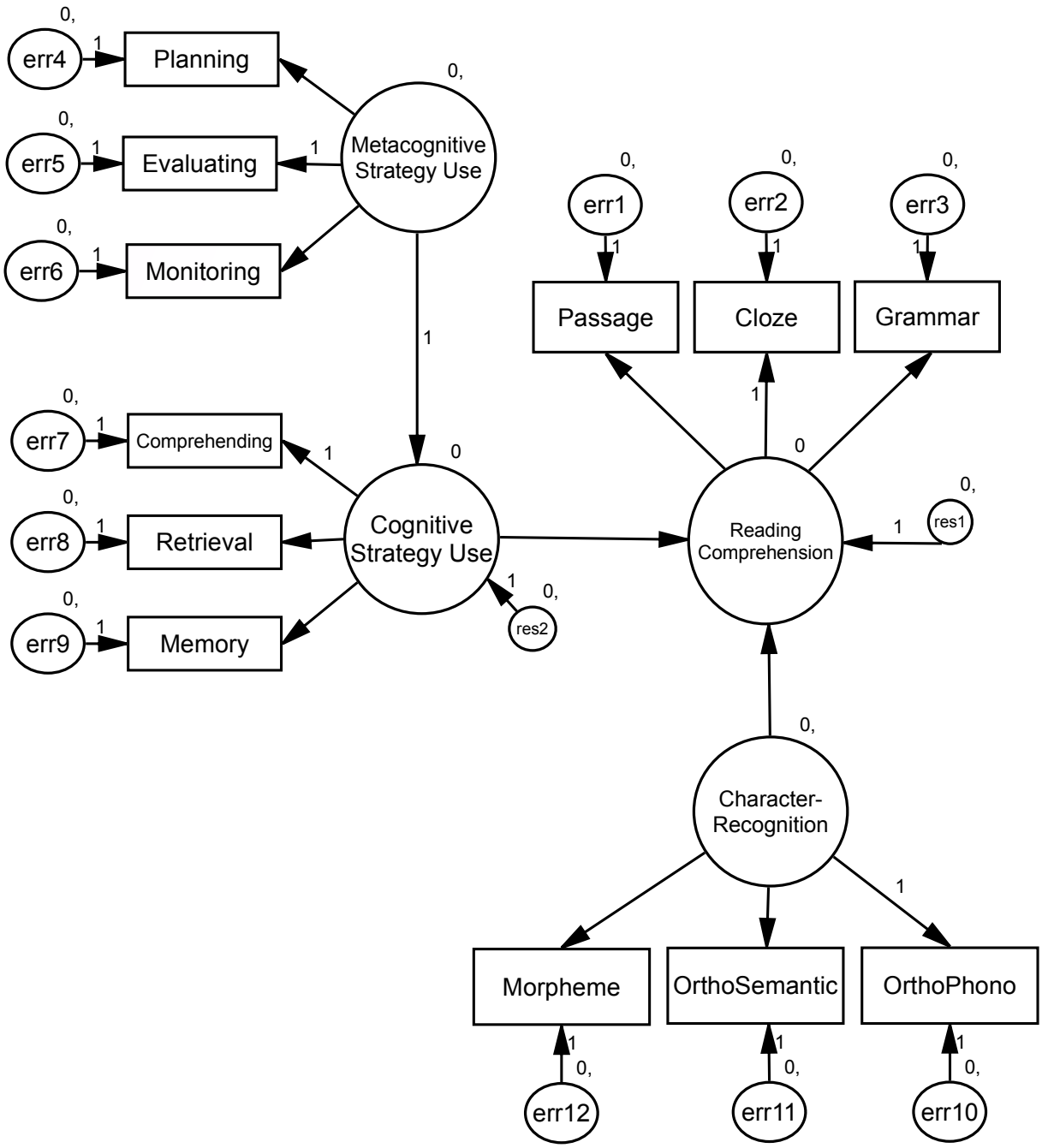


Figure 1. Proposed model of reading Chinese as an L2

This dissertation's preliminary model of L2 Chinese reading, shown in Figure 1, is centered around reading comprehension ability and indicates hypothesized relationships between such ability and measures of a reader's character-recognition ability and strategy use. The language-background element is missing from this model, as the model itself is nested in language background. Likewise, group differences in character-recognition skills, strategy use, and reading-test performance will be examined before the hypothesized model of L2 Chinese reading. The participants in this research are CFLLs, Man-HLLs, or Can-HLLs.

2.7 Research Questions

1. To what extent can language background explain the variance in performance on the reading comprehension test, the character-recognition test, and strategy-use survey?
2. Across the three groups of participants, what are the relationships between test performance of the reading comprehension test, character-recognition test-taker perception of strategy-use while taking the reading test?
3. Across the three groups of participants, to what extent do character-recognition versus strategy use factors affect test performance of the reading test?

CHAPTER 3

METHODOLOGY

3.1 Description of the Research Phases

After the instruments had been developed, two pilot studies were conducted followed by the main study. The first pilot study was designed to ensure the answer keys were acceptable for native speakers. The second pilot study was to evaluate the item difficulties for the target population and to delete items which were too difficult or too easy. After the measurement instruments were calibrated, recruiting emails were sent out for the main data collection. Recruiting emails were sent out to program administrators and instructors whose Chinese language programs offer intermediate and advanced language courses, such as the second and third year Chinese at the university level. Email recipients were asked to distribute recruitment flyers to their students who might meet the recruiting criteria. Chinese language learners who were interested in participating in the study could access the tests on an online platform. The two recruiting steps were implemented to decrease the chance of unqualified participants taking the test, and especially to avoid robotic responses developed for collecting monetary compensation through participating in online surveys. In the subsequent data analysis stage, profile analysis was implemented to answer the first research question. This study also utilized SEM to evaluate the hypothesized model shown in Figure 1, as a means of answering the second and third research questions.

3.2 Data-Collection Procedures

As previously mentioned, data were collected in two phases: two small-scale pilot sessions and a larger-scale data collection phase. Descriptions of data-collection procedures and the methods of analysis adopted in each phase are provided below.

3.2.1 Small-scale pilot studies. Pilot sessions were conducted with two samples: one consisting of five highly educated native speakers of Mandarin, and the other of 11 Chinese language learners with intermediate or advanced Chinese reading levels. In the study, the intermediate level was defined as being able to read basic narrative and descriptive passages related to personal and social topics, and the advanced level was defined as being able to read conventional narrative and descriptive passages and being able to identify the main ideas, facts and supporting details of the texts (Swender, Conrad, & Vicars, 2012). A concise version of the test, containing only the reading test and the Chinese character test, was administered to the pilot group of native speakers. In this phase, the reading test consisted of 20 multiple-choice items focused on grammar, a cloze subtest with two passages, containing 20 blanks, and a passage-comprehension test with five passages, each with four multiple-choice items. In the cloze subtest, each blank contained only one character. Seven characters separated each blank. The Chinese character test had 10 multiple-choice items for each of the three subtests. This first pilot session was carried out in May 2015.

In the first pilot session, to ensure the quality of the test, the multiple-choice items that did not receive the same responses from all five native speakers were revised and re-administered to the same native-speaking participants. This revision process continued until all five participants chose the same option. However, the fill-in-the-blank cloze test underwent different revision procedures. All the responses provided by the five native speakers were regarded as correct answers for purposes of the main study.

After revising the test items, convenience sampling was used to recruit 11 Chinese language learners with target reading levels; this group took the complete test with all the four parts of the study. The responses to the reading test and the character-recognition test were

analyzed further. Items whose facilities were higher than 0.8 (indicating that they were too easy for the target participants) or lower than 0.3 (indicating that they were too difficult) were excluded from the final version of the measurement instruments. The responses to the language-background and strategy-use surveys were glanced over and no unusual responses were spotted. Reliability coefficients, calculated as Cronbach's alpha, of the reading test, the character-recognition test, and the six strategy types were summarized in Table 1. Considering the small numbers of test takers and items, some reliability coefficients are low, especially the alarmingly low reliability of memory strategies, which was zero. Memory strategies in the survey include interpreting hidden ideas, translating, summarizing the main information, and rereading the text. The zero reliability may reflect the unclear reading goals and purposes for the participants, leading to inconsistent use of memory strategies. However, considering the empirical support of memory strategies found in the studies of Phakiti (2008) and Purpura (1997), this strategy type was remained in the survey.

Table 1

Reliability Coefficients of the Second Pilot Session

$N = 11$	k	Reliability
Grammar	8	0.49
Cloze	12	0.86
Passage	10	0.78
Comprehending	4	0.92
Memory	4	0.00
Retrieval	5	0.38
Planning	6	0.76
Monitoring	5	0.63
Evaluating	6	0.61
Ortho-Phono	4	0.91
Ortho-Semantic	4	0.57
Morpheme	5	0.64

3.2.2 Main data collection. After the measurement instruments were finalized, recruiting emails were sent out for the main data collection phase from September 2015 to March 2016. Recruiting emails were sent out to program administrators and instructors whose Chinese language programs offer classes at the target proficiency levels, such as the second year Chinese. In the emails, recipients were asked to distribute recruitment flyers to their students who might meet the recruiting criteria; and interested students could access the online tests through the url link listed in the recruitment flyer. To participate in the main study, participants had to be at least 16 years old, have English as their dominant language, and be able to read a passage containing more than 250 Chinese characters (either in Simplified or Traditional characters). Additionally, participants were solicited to be members of one of three groups.

1. CFLLs: participants who were native speakers of English, and both of whose parents were native speakers of English.
2. Man-HLLs: participants whose strongest language was English, and both of whose parents were native speakers of Mandarin.
3. Can-HLLs: participants whose strongest language was English, and both of whose parents were native speakers of Cantonese.

Recruiting emails were sent through the Chinese Language Teachers Association (CLTA) in the United States, the American Council on the Teaching of Foreign Languages (ACTFL), and the Canadian Association of Applied Linguistics, as well as over 30 individual Chinese-language programs in California, Florida, Georgia, Hawaii, Illinois, Kansas, Maryland, New Jersey, New York, Ohio, Oregon, Rhode Island, Texas, Vermont, Utah, British Columbia in Canada, and Singapore.

3.3 Measurement Instruments

Before taking the test, each participant was shown a consent form for this study, and was notified that checking the "Submit" box at the end of the test indicated their agreement to participate in the study. The online session was divided into four sections: a language-background survey, a reading test, a strategy-use questionnaire, and a Chinese-character recognition test. The four sections were presented to all participants in the particular order, shown in Table 2 (see Appendix A for the complete test).

Table 2

Structure of Measurement Instruments

Test	Subtest
Language Background Survey	
Reading Test	Grammar
	Cloze
	Passage-Comprehension
Strategy Use Survey	
Chinese-Character Test	Ortho-Phonological
	Ortho-Semantic
	Morpheme

3.3.1 Language-background survey. The language-background survey consisted of questions regarding the participant's age; birthplace; years of learning Chinese and the type(s) of setting in which it was learned; years lived in a Chinese-speaking country; dominant language(s) and the dominant language(s) of her/his parents; amount of Mandarin use at home; self-assessment of their proficiency in listening, speaking, reading and writing, and experience of learning languages other than her/his own.

3.3.2 Reading test. This test consisted of three subtests. First, multiple-choice items targeted learners' Chinese grammar; second, a fill-in-blank cloze subtest targeted their general

linguistic knowledge of Chinese; and third, a passage-comprehension subtest targeted their reading comprehension skills. After revising the measurement instruments based on the results of the second pilot study, the grammar subtest had eight items, the cloze subtest had 12, and the passage-comprehension subtest, 10.

The grammar rules examined in the grammar subtest were selected from levels four to six of the 2013 *Hanyu Shuiping Kaoshi* (Chinese Proficiency Test) (Hanban/Confucius Institute, n.d., retrieved on January 2, 2015); two mock Test of Proficiency-*Huayu* (TOP) for advanced and superior reading (Steering Committee for the Test Of Proficiency-*Huayu*, n.d., retrieved on January 2, 2015). Although the grammar rules were chosen from the exams, the prompts and options were newly designed for the current study by the researcher. All questions were presented with both simplified and traditional characters so test takers were able to read the version they were most familiar with. The grammar subtest was designed to measure the participants' use of grammatical rules/constraints and phrasal verbs, as shown in Table 3.

Table 3

Types of Questions Asked in the Grammar Subtest

Question	Examples
Grammatical rules and constraints	比 in basic comparison construction 把 as a classifier for combs 吗 as a question marker for yes/no questions 对 to mean "regarding, with regards to" verb verb 看 to indicate that an action happens briefly, e. g. 穿穿看 (*chuānchuān kàn* "put on the cloth for a short while") verb + 好 to indicate an action is completed
Phrases	除非…否则 "unless" 不是…而是 "rather than"

The passages in the passage-comprehension subtests were selected from levels four to six of the 2013 *Hanyu Shuiping Kaoshi* (Chinese Proficiency Test) (Hanban/Confucius Institute, n.d., retrieved on January 2, 2015); a mock Test of Proficiency-*Huayu* (TOP) for advanced reading (Steering Committee for the Test Of Proficiency-*Huayu*, n.d., retrieved on January 2, 2015); the 2007 Advanced Placement Chinese exam (College Board Advanced Placement Program, 2009); and a news story from People's Daily (Ye & Dong, 2011). However, the reading-comprehension questions were newly created by the researcher for use in this study. All questions were presented with both simplified and traditional characters.

The passage-comprehension subtests consisted of five passages, ranging from 340 to 430 characters. Their topics included an exchange of email with parents, news articles about Chinese college education and job-searching by overseas returnees to China, Chinese mythology, and scientific exploration in Antarctica. The comprehension questions were designed to examine four main comprehension skills: scanning for explicit information, skimming for gist, connecting explicit information, and drawing inferences. Table 4 summarizes the types of questions in this subtest.

The two passages in the cloze part of the test were adopted from digital, interactive material developed by Fleming, Hiple, and Ning (2002) for use as a textbook for a third-year online course in Chinese reading and writing at an American university. Both passages contained between 120 and 140 Chinese characters; one was about a scenic attraction in China, and the other about highway systems. This subtest was included to measure participants' ability to comprehend texts drawing on their vocabulary and grammar knowledge.[6]

[6] When the cloze subtest was first developed, the blanks targeted knowledge of vocabulary, grammar, pragmatics, and discourse. However, after deleting blanks with item difficulties that were too high and too low, only blanks targeting vocabulary and grammar knowledge remained.

Table 4

Types of Questions Asked in the Passage-Comprehension Subtest

Question type	Explanation	Example
Skimming for gist	These questions asked participants to identify answers that correctly summarize the main idea of the whole passage.	Which of the following best describes the overall goal of the plan discussed in the article?
Scanning for explicit information	These questions asked participants to identify answers about specific details mentioned in the passage.	According to the article, how do Chinese college students differ from college students in Europe and the United States?
Connecting explicit information	These questions asked participants to identify answers by connecting explicit details in the passages.	What does Peter usually do on Wednesday afternoon?
Drawing inferences	These questions asked participants to identify answers that were not specified in the passages, based on other details that were.	What can we know about 极昼?

3.3.3 Strategy-use survey. The strategy-use survey was adopted from a previous study (Phakiti, 2008) with minor changes matching the features of Chinese reading. The survey was designed to gain an understanding of the strategies the participants used while they were taking the reading test. There were 30 statements about such strategies, which the participants rated on a 4-point Likert scale, with the following values: “Strongly agree,” “Agree,” “Disagree,” and “Strongly disagree.” Degrees of dis/agreement of certain strategy-use was adopted because Tseng, Dörnyei, and Schmitt (2006) argue that the perceived strategy-use frequency cannot be viewed as an interval variable and a dis/agreement scale tapping into personal attitude may be more appropriate for Likert-scale strategy-use surveys. Test takers were asked to answer the strategy-use survey immediately after the reading test, to help ensure that their responses

reflected their online strategy use when taking the reading subtests rather than their understanding of their own strategy use for general language learning contexts. Developed based on information-processing theories, the cognitive and metacognitive strategies were both included in the strategy-use survey, with the former type including strategies aimed at *comprehension*, *memorization* of the text, and *retrieving* prior knowledge, while the metacognitive type consisted of strategies used for *planning*, *monitoring*, and *evaluating* reading comprehension (Phakiti, 2008). Table 5 presents the strategy composites used in the survey.

Table 5

Taxonomy of the Cognitive and Metacognitive Strategies Covered by the Strategy-Use Survey

Category	Strategy	*k*	Item Nos.
Cognitive	Comprehending	4	7, 8, 9, 10
	Memory	4	11, 12, 13, 15
	Retrieval	5	14, 16, 17, 18, 19
Metacognitive	Planning	6	1, 2, 3, 4, 5, 6
	Monitoring	5	20, 21, 24, 29, 30
	Evaluating	6	22, 23, 25, 26, 27, 28
Total		30	

The comprehending strategies include cognitive activities related to understanding the content and the relationship between ideas in the passage, predicting the following content before reading, and analyzing the author's intention. The memory strategies are related to interpreting

the hidden ideas, translating the text to L1, summarizing the main ideas, and re-reading the passage when inadequate comprehension occurs. The retrieval strategies are related to relating the passage to prior knowledge, identifying the importance of the information, guess meaning of unfamiliar words based on context clues and Chinese-character knowledge, and applying learned grammar rules.

The planning strategies are related to clarifying and planning steps before reading, setting reading goals, being aware of how well the original plans are accomplished after reading, and skimming thorough comprehension questions before answering them. The monitoring strategies are related to time management, monitoring the progress of taking the reading test, and monitoring one's own concentration, affective, and comprehension status. The evaluating strategies are related to evaluating comprehension quality, test performance, test progress, original reading plans, and reading goals, adjust reading speed, and test-taking speed, and correct misunderstanding.

3.3.4 Chinese-character test. The Chinese character test consisted of ortho-phonological, ortho-semantic, and morpheme subtests. In the ortho-phonological subtest, a pseudo-character made up of two real characters was presented at the beginning of each item, and the participants were asked to choose the option with the closet pronunciation to it, from among another the three options. A correct answer was the phonetic component of the pseudo-character and a distractor was the semantic radical of the pseudo-character, and another distractor was a character with a similar shape to the target pseudo-character.

In the ortho-semantic subtest, a drawing of something that does not exist in the real world was presented at the beginning of each item, and the participants were asked to choose the option that best described it. Three types of drawing were included: insects, wooden objects, and

luminous or flaming objects, representing three semantic radicals. Each item included three options, each of which was a pseudo-character made up of two real characters. To successfully complete this task, a participant would need to know the semantic radicals representing the drawing and their positions as semantic radicals. Again, two distractors were presented in the options: 1) a pseudo-character with the same sub-character components as the correct answer, but violating the rules of positional regularity; and 2) a pseudo-character containing the same character as the phonetic component of the correct answer.

The morpheme subtest was designed to assess the participants' knowledge of characters' multiple meanings. Each item consisted of three words that shared one character, and this target character was located in the same position within the three options. The participants were asked to identify the option whose shared character had the meaning that differed the most from those of the other two options. For example, the three words in one item were 开始 (*kāishǐ*, "to start"), 开学 (*kāixué*, "school opening"), and 开车 (*kāichē*, "to drive"). The correct answer was 开车 (*kāichē*, "to drive"), because the 开 in this word does not mean "to start" as it does in the other two words.

3.4 Participants

This section presents the participants' responses to the language-background survey, and the following section summarizes their responses to the reading test, the strategy-use survey, and the character-recognition test. As I was examining the responses, it became clear from the descriptive statistics that the participants who identified their parents as native speakers of English, should be further subdivided into Singaporean Chinese Mother-Tongue Language Learners (Singaporean CMTLLs) and CFLLs even though the study design originally included

only the three participant groups mentioned in the previous section. In Singapore, although English is the official language for all Singaporeans, there are three other "mother tongue" languages for students to learn at school (Ministry of Education, Singapore, 2016a, 2016b). In other words, the term "mother tongue" in the Singaporean setting does not necessarily indicate the participants' L1 or home language. On the other hand, considering the large population of ethnic Chinese living in Singapore (Department of Statistics Singapore 2015), it would be safe to assume that Mandarin Chinese would be accessible outside of language classes and for the Singaporean participants who identified their parents as native speakers of English, Mandarin Chinese in Singapore would be different from a foreign language. As such, the term Singaporean CMTLLs is used to reflect the unique multi-lingual context of Singapore even though they identified their parents as native speakers of English. After data screening, 85 participants were left in the sample. The descriptive statistics of their language background are summarized in Tables 6 to 9.

Table 6

Summary of Participants' Birthplaces

Birthplace	*n*	Canada	China	Germany	Hong Kong	Singapore	Taiwan	USA	NA
Singaporean CMTLLs	14	0%	0%	0%	0%	100%	0%	0%	0%
CFLLs	19	0%	0%	5%	0%	0%	0%	89%	5%
Man-HLLs	38	8%	8%	0%	0%	5%	29%	50%	0%
Can-HLLs	14	7%	7%	0%	7%	7%	0%	71%	0%
Total	85	5%	5%	1%	1%	20%	13%	54%	1%

Table 7

Previous Mandarin Learning Experience, in Years

	n	Age *M (SD)*	AoA[a] *M (SD)*	Chinese Weekend School *M (SD)*	Secondary School *M (SD)*	College *M (SD)*
Singaporean CMTLLs	14	18.00 (0.85)	0.00 (0.00)	6.31 (6.17)	4.93 (0.26)	0.25 (0.60)
CFLLs	19	21.85 (8.02)	0.06 (0.34)	0.83 (1.01)	1.63 (1.45)	1.76 (1.86)
Man-HLLs	38	24.03 (7.19)	1.95 (3.37)	3.63 (3.96)	1.64 (2.27)	0.41 (0.89)
Can-HLLs	14	22.93 (4.95)	0.36 (0.89)	1.54 (3.46)	1.39 (1.63)	1.54 (1.66)

a. AoA: age of arrival in an English-speaking country

As Table 6 indicates, just over half the participants were born in the United States (*n* = 46, 54%). Among these American participants, Man-HLLs made up the largest group, at 19 or 41.3%; CFLLs the second largest, at 17 or 37%; and Can-HLLs the smallest, at 10 participants or 21.7% of the American total. In contrast to American participants, most Singaporean participants were CFLLs, which might at first seem counter-intuitive, given the substantial population of ethnic Chinese there. However, English has been used as a *lingua franca* in Singapore since the era of British rule, so it is understandable that most Singaporean participants identified their parents as native speakers of English. Among those participants born in China and Taiwan, all but one were identified as Man-HLLs, and most of Can-HLLs (*n* = 10, 71%) were born in the States.

As shown in Table 7, the four groups' average ages ranged from 18.00 to 24.03 years. Among those who were in China, German, and Taiwan, most moved to an English-speaking country when aged younger than one year. However, the mean and standard deviation of Man-HLLs age of arrival in an English-speaking country (AoA) were slightly higher than those of the other three groups. In general, Singaporean CMTLLs had spent the longest studying Chinese in weekend schools and secondary schools, and Man-HLLs the second longest. CFLLs who were not from Singapore (CFLLs) had spent the least time studying in a Chinese weekend school, in terms of mean and standard deviation. Although Can-HLLs seemed to have studied Chinese longer than CFLLs had, the difference was not substantial.

With regard to the time spent studying Chinese in secondary school, Singaporean CMTLLs reported the longest duration. The likely reason for this is that Singaporeans take another language in addition to English, as required by the Mother Tongue Language (MTL) policy (Ministry of Education, Singapore, 2016a). This "requires all students … to study their

respective official MTL: Chinese, Malay and Tamil" in elementary school onward (Ministry of Education, Singapore, 2016b). Table 8 summarizes participants' Mandarin use at home. It shows that Singaporean CMTLLs and Man-HLLs reported more Mandarin use with their parents than with their grandparents, and more than 30% of Man-HLLs reported that they always used Mandarin with their grandparents and other relatives. For all four groups of participants, Mandarin use with siblings was lower than with other family members, and between 43% and 79% of the participants in the four groups reported that they did not speak Mandarin with their siblings at all.

Table 8

Summary of Mandarin Use at Home

		Mandarin use at home					
		w/ Mother	w/ Father	w/ Siblings	w/ Maternal grand-parents	w/ Paternal grand-parents	w/ Other relatives
Singaporean CMTLLs (*n* = 14)	Always	0%	7%	0%	43%	7%	0%
	Very Often	0%	0%	0%	7%	21%	14%
	Often	21%	21%	0%	21%	14%	14%
	Sometimes	43%	43%	50%	7%	21%	50%
	Not at all	36%	29%	43%	14%	29%	21%
	Missing	0%	0%	7%	7%	7%	0%
CFLLs (*n* = 19)	Always	0%	0%	0%	0%	0%	0%
	Very Often	5%	0%	0%	5%	0%	0%
	Often	0%	0%	0%	0%	0%	0%
	Sometimes	11%	16%	26%	0%	0%	11%
	Not at all	84%	84%	74%	95%	100%	84%
	Missing	0%	0%	0%	0%	0%	5%
Man-HLLs (*n* = 38)	Always	34%	34%	5%	66%	63%	53%
	Very Often	26%	24%	8%	11%	11%	11%
	Often	18%	21%	21%	11%	8%	13%
	Sometimes	18%	16%	21%	3%	3%	11%
	Not at all	3%	5%	42%	8%	13%	5%
	Missing	0%	0%	3%	3%	3%	8%
Can-HLLs (*n* = 14)	Always	0%	0%	0%	7%	0%	7%
	Very Often	0%	0%	0%	0%	7%	0%
	Often	7%	14%	0%	0%	0%	0%
	Sometimes	36%	21%	21%	0%	0%	21%
	Not at all	57%	64%	79%	86%	86%	64%
	Missing	0%	0%	0%	7%	7%	7%

In sum, of the four groups, Man-HLLs reported the highest level of Mandarin use at home, and more than a third of them stated that they always spoke Mandarin with their parents. In addition, despite compulsory MTL education, and having spent the longest studying Mandarin of any of the four groups, about a third of Singaporean CMTLLs reported that they did not use Mandarin with their parents at all, and about 40% of them did so only sometimes. Unsurprisingly, less than one-fifth of CFLLs used Mandarin with their parents, while Can-HLLs demonstrated the lowest overall level of Mandarin use at home.

Although self-reported proficiency is not an accurate indicator of actual proficiency, it is useful in understanding participants' confidence in their language skills. Self-reported Mandarin proficiency is therefore included in Table 9. Aligned with their Mandarin use at home, Man-HLLs' assessment of their proficiency was the highest, with Singaporean CMTLLs second; CFLLs and Can-HLLs reported similar self-assessed proficiencies summarized in Table 9. When comparing the four language skills across the four participant groups, listening was scored highest, followed by speaking, reading, and writing.

In Table 9, among these four skills, the low self-assessment of Chinese reading may relate to the perceived challenge of learning to reading Chinese texts. For example, only around 60% of Singaporean CMTLLs rated their Chinese reading as proficient, and none rated it as exemplary, despite all of them having had more than 10 years of Mandarin learning. Similarly, less than half of Man-HLLs evaluated themselves as proficient or exemplary Chinese readers. In comparison to Singaporean CMTLLs and Man-HLLs, CFLLs and Can-HLLs' self-assessment of their Chinese reading ability was even lower: with around 11% of CFLLs and no Can-HLLs rating themselves as proficient.

Table 9

Summary of Self-Report Mandarin Proficiency

		Mandarin proficiency in			
		Listening	Speaking	Reading	Writing
Singaporean CMTLLs (*n* = 14)	Exemplary	0%	0%	0%	0%
	Proficient	79%	50%	57%	43%
	Developing	21%	50%	29%	43%
	Minimal	0%	0%	14%	14%
	None	0%	0%	0%	0%
CFLLs (*n* = 19)	Exemplary	0%	0%	0%	0%
	Proficient	16%	16%	11%	11%
	Developing	68%	68%	79%	63%
	Minimal	11%	11%	11%	26%
	None	5%	5%	0%	0%
Man-HLLs (*n* = 38)	Exemplary	32%	29%	18%	11%
	Proficient	53%	47%	29%	18%
	Developing	16%	24%	37%	47%
	Minimal	0%	0%	16%	13%
	None	0%	0%	0%	11%
Can-HLLs (*n* = 14)	Exemplary	0%	0%	0%	0%
	Proficient	14%	14%	0%	7%
	Developing	79%	64%	71%	29%
	Minimal	7%	21%	29%	64%
	None	0%	0%	0%	0%

3.5 Subtest Scores of the Four Groups

The boxplots of subtest scores for the four groups of participants are presented in Figures 2, 3, and 4, with Figure 2 covering the three reading subtests, Figure 3 the six strategy types, and Figure 4 the three character-recognition subtests. The corresponding descriptive statistics appear in Tables 10, 11 and 12.

3.5.1 Subtest scores in the reading test. As shown in Figure 2, the Singaporean CFLL group included two extreme values (relative to other members of the same group), but these two values fell in the range of three standard deviations from the mean of the 85 participants. Also, the grammar distribution for Singaporean CMTLLs is not visible because of its minimal standard deviation (0.07). Singaporean CMTLLs had the highest means among the four groups and across

the three subtests; they also had the smallest standard deviations, indicating that they gave more uniform responses on each subtest in comparison to the other three groups. Man-HLLs had the second-highest subtest means, but their standard deviations showed a wider distribution than those of Singaporean CMTLLs and CFLLs. Among the three subtests, CFLLs scored higher in the grammar and passage-comprehension subtests than the cloze test. Although the cloze subtest was the most difficult one for the four groups, CFLLs' scores were the lowest and most positively skewed of all four groups' distributions.

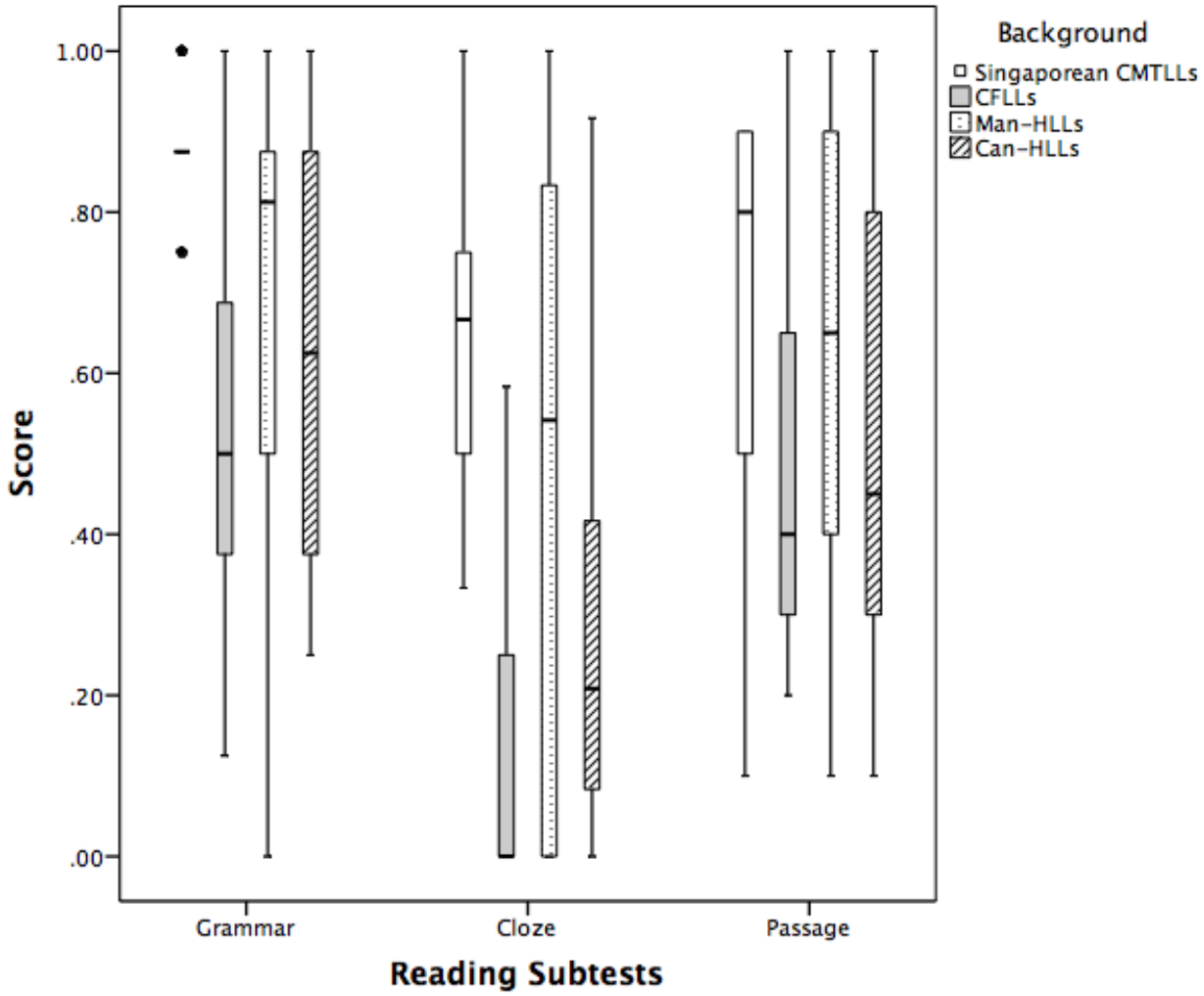


Figure 2. Boxplots of the reading subtests

Table 10

Descriptive Statistics for the Reading Subtests

Range: 0-1	Grammar *M (SD)*	Cloze *M (SD)*	Passage *M (SD)*
Singaporean CMTLLs	0.88 (0.07)	0.66 (0.17)	0.67 (0.24)
CFLLs	0.53 (0.23)	0.18 (0.30)	0.48 (0.25)
Man-HLLs	0.71 (0.27)	0.50 (0.37)	0.61 (0.28)
Can-HLLs	0.60 (0.24)	0.31 (0.28)	0.53 (0.31)

3.5.2 Subtest scores in the strategy-use survey. As can be seen from Figure 3 and Table 11, the distributions of self-reported strategy-use of the six strategy types were narrower (in comparison to the group distributions of the reading subtests). In terms of group means, comprehending, retrieval, and monitoring strategies were the strategies most commonly reported by Singaporean CMTLLs, CFLLs, and Man-HLLs. Among the four groups, Can-HLLs reported the least strategy use – although such a difference may have been minimal and statistically nonsignificant. It should be noted that in Figure 3, the asterisks indicate the extreme values within each group, but not within the sample as a whole.

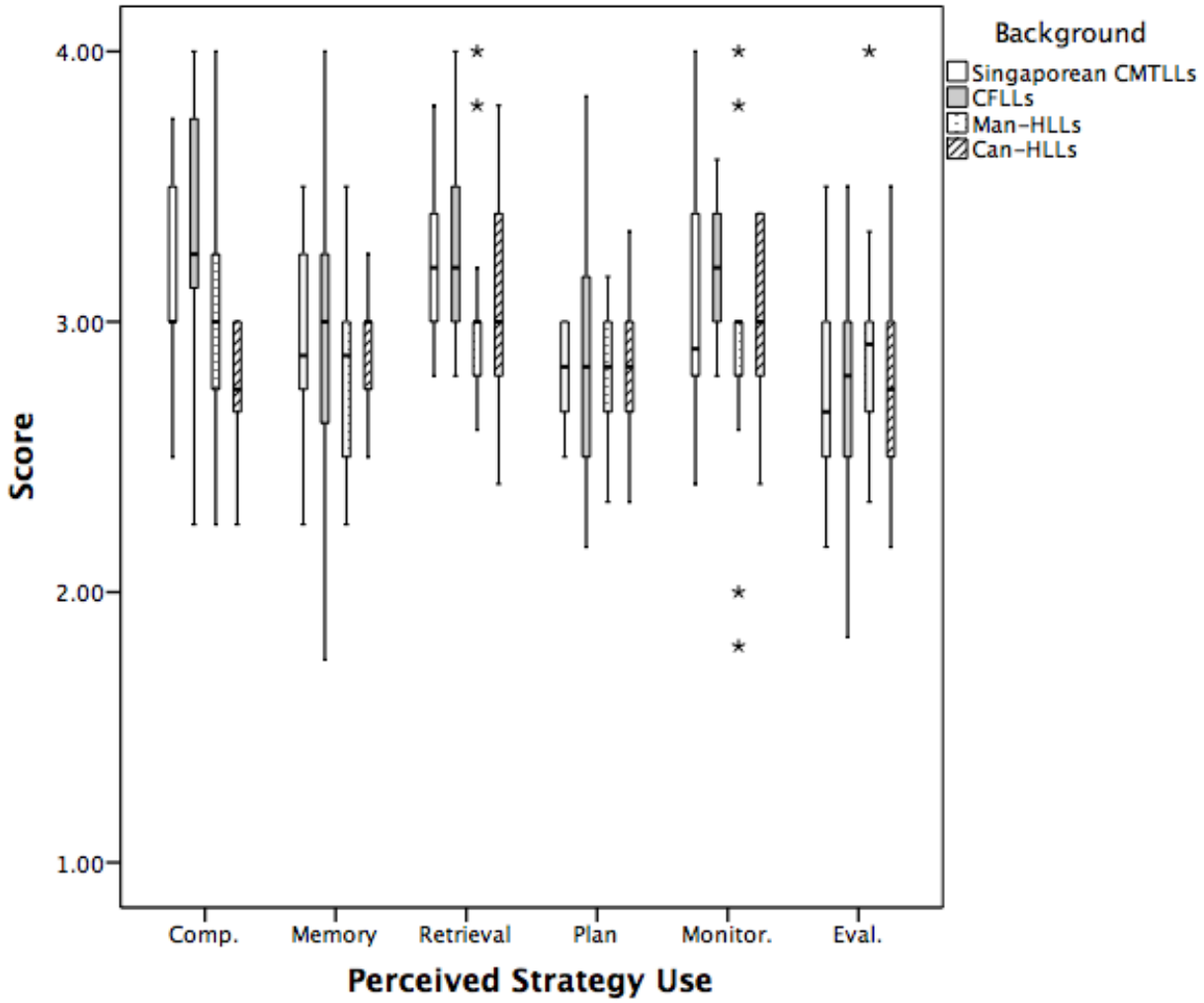


Figure 3. Boxplots of strategy use. Comp. = comprehension strategies, Plan. = planning strategies, Monitor. = monitoring strategies, Eval. = evaluating strategies.

Table 11

Descriptive Statistics for Strategy-Use Survey

Range: 1-4	Comprehending *M (SD)*	Memory *M (SD)*	Retrieval *M (SD)*	Planning *M (SD)*	Monitoring *M (SD)*	Evaluating *M (SD)*
Singaporean CMTLLs	3.14 (0.34)	2.91 (0.32)	3.23 (0.29)	2.75 (0.35)	3.01 (0.42)	2.79 (0.19)
CFLLs	3.34 (0.50)	2.92 (0.55)	3.26 (0.36)	2.81 (0.58)	3.15 (0.40)	2.89 (0.43)
Man-HLLs	3.07 (0.46)	2.81 (0.37)	3.03 (0.31)	2.87 (0.40)	2.94 (0.44)	2.85 (0.36)
Can-HLLs	2.82 (0.29)	2.95 (0.29)	3.03 (0.38)	2.79 (0.35)	3.01 (0.32)	2.81 (0.40)

3.5.3 Subtest scores in the character-recognition test. Figure 4 and Table 12 summarize the scores of the three character-recognition subtests for the four participant groups. In terms of the means of the ortho-phonological and ortho-semantic subtests, Singaporean CMTLLs did not show any clear advantage over the other three groups, as they had in the reading subtests. In the ortho-phonological subtest, while all four groups had similar means, Singaporean CMTLLs and CFLLs showed narrower distributions than Man-HLLs and Can-HLLs did. In the ortho-semantic subtest, although Singaporean CMTLLs had a slightly higher mean than the other three groups, the differences between Singaporean CMTLLs and the other three groups may not be statistically significant. Moreover, Man-HLLs' boxplot demonstrated the narrowest distribution among the four groups, which is less clear based on its descriptive statistics. In the morpheme subtest, Singaporean CMTLLs had the highest mean and smallest standard deviation, and Man-HLLs had a wider distribution than Singaporean CMTLLs did although Man-HLLs' mean and standard deviation were similar to those of Singaporean CMTLLs.

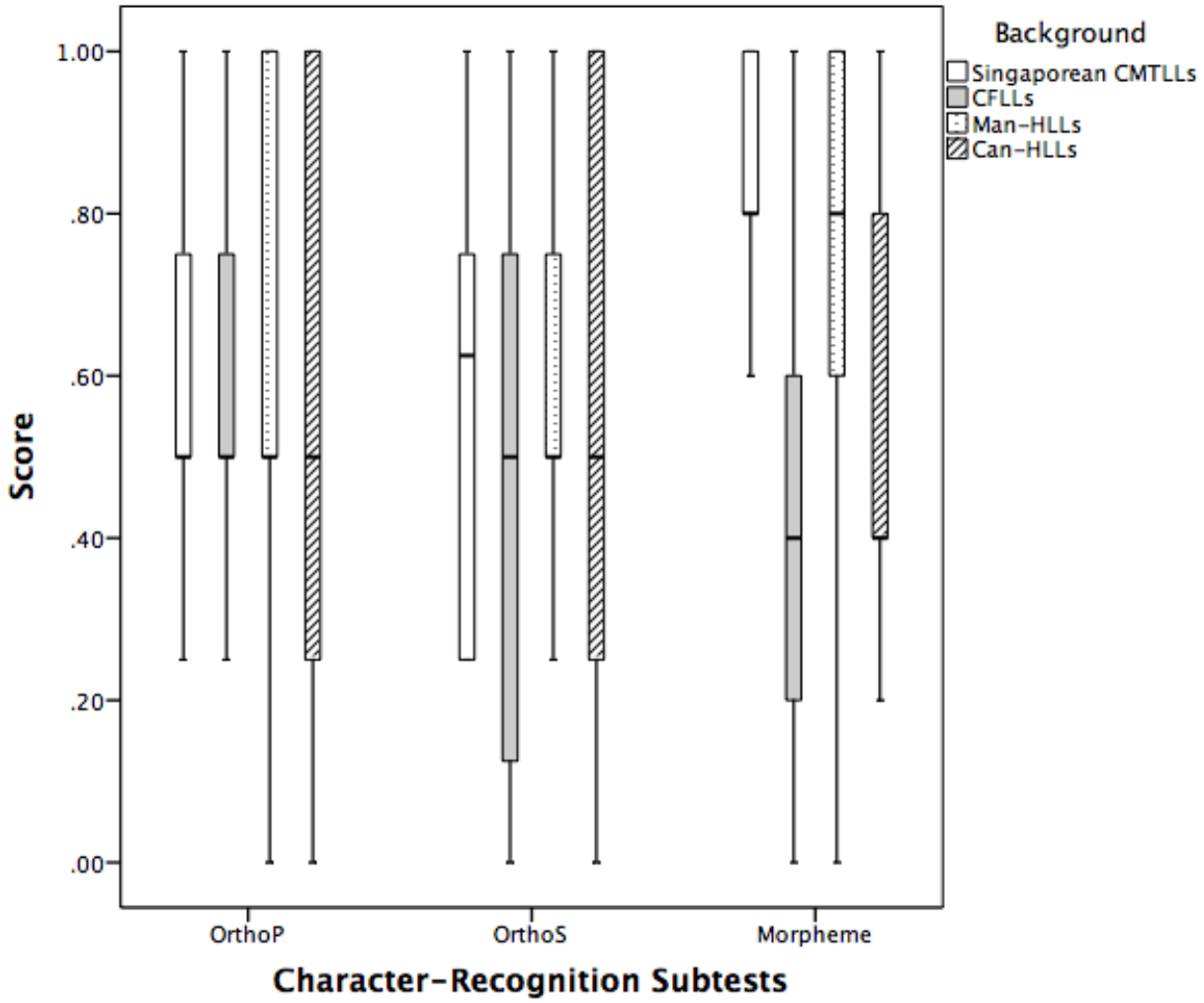


Figure 4. Boxplots of the character-recognition subtests. OrthoP = ortho-phonological subtest, OrthoS = ortho-semantic subtest.

Table 12

Descriptive Statistics for the Character-Recognition Subtests

Range: 0-1	Ortho-Phono *M (SD)*	Ortho-Semantic *M (SD)*	Morpheme *M (SD)*
Singaporean CMTLLs	0.57 (0.20)	0.61 (0.28)	0.81 (0.18)
CFLLs	0.57 (0.29)	0.47 (0.36)	0.44 (0.26)
Man-HLLs	0.61 (0.34)	0.59 (0.31)	0.74 (0.26)
Can-HLLs	0.55 (0.39)	0.52 (0.36)	0.54 (0.28)

In sum, in terms of test performance across the 12 subtests, Singaporean CMTLLs had higher mean scores in the reading subtests and the morpheme subtest than the other three groups did; but they had no such advantage in the strategy-use and character-recognition subtests. In general, the four groups reported similar use of the six strategies and performed similarly on the two ortho-subtests. Based on the above descriptive statistics and boxplots, it can be seen that Singaporean CMTLLs had considerable similarity to Man-HLLs as language learners, despite having stated that both of their parents were native speakers of English. Such similarity may result from Singaporean CMTLLs' greater access to Mandarin than typical CFLLs (or even HLLs) have in the United States and Canada.

3.6 Data Analysis for the Main Study

This section first reviews the present study's data-analysis and data-screening procedures, then presents the results of reliability analysis of the subtests, and concludes with explanations of the statistical methods implemented to answer the three research questions.

3.6.1 Data-analysis procedures. The data were first screened to ensure that they met the assumptions of the chosen statistical analyses. Although the main goal of the study was to examine common constructs of L2 Chinese reading across language learners' various language backgrounds, the 12 variables were first divided into the reading test, strategy-use survey, and the character-recognition test, and then the extent of group differences were examined with three profile analyses to answer the first research question. Given that multiple comparisons were made with a relatively small sample (n = 85), the Results and Discussion chapters that follow focus on interpreting effect sizes caused by group differences, rather than on significance values (p value).

After answering the first research question, single-group SEM was adopted to answer the second and the third research questions, using the hypothesized model shown in Figure 1 presented in Chapter 2. This model conceptualizes the character-recognition skills as a construct of 1) lower-level processing skills, 2) the comprehension skills examined in common reading tests as a construct of higher-level processing skills, and 3) strategy-use as another construct of higher-level processing skills.

3.6.2 Data screening. At the end of the main data-collection period, 132 participants had taken the test, but 90 participants were left after data screening. First, the researcher checked submissions' IP addresses and responses, in order to delete any multiple submissions made by the same individual, as well as submissions by participants who were unlikely to be English-dominant and/or those who reported no proficiency in Chinese reading. Specifically, the deletion criteria were 1) responses that emanated from the same IP addresses; 2) participants who said they were not born in an English-speaking country but had moved to an English-speaking country as adults for less than 10 years, i.e., after 2006; 3) participants who reported having no proficiency in Chinese reading; and 4) participants who failed to answer a substantial number of questions in the language-background survey, the reading test, or the strategy-use survey.

Next, the data-screening procedures recommended by Tabachnick and Fidell (2012) were followed, as summarized in Table 13. Univariate descriptive statistics were checked and no out-of-range values were found; all means and standard deviations were plausible. The four univariate outliers were detected and shown in Table 13, and these outliers were all in the strategy-use survey which may be related to the smaller standard deviations in the six strategy variables.

Table 13

Checklist for Screening Data (cited from Tabachnick & Fidell, 2012, p. 91)

1		Inspect univariate descriptive statistics for accuracy of input
	a	Out-of-range values
	b	Plausible means and standard deviations
	c	Univariate outliers
2		Evaluate amount and distribution of missing data: deal with problem
3		Check pairwise plots for nonlinearity and heteroscedasticity
4		Identify and deal with non-normal variables and univariate outliers
	a	Check skewness and kurtosis, probability plots
	b	Transform variables (if desirable)
	c	Check results of transformation
5		Identify and deal with multivariate outliers
	a	Variables causing multivariate outliers
	b	Description of multivatiate outliers
6		Evaluate variables for multicollinearity and singularity

Table 14

Mean Scores of the Univariate and Multivariate Outliers

Back-ground	Birth place	Age	AoA[a]	Reading Test			Strategy-Use Survey						Chinese-Character Test		
				Grammar	Cloze	Passage	Comp.[b]	Memory	Retrieval	Planning	Monitor.[c]	Eval.[d]	OrthoP[e]	OrthoS[f]	Morpheme
Univariate outliers[g]															
Man-HLL	USA	24	0	0.75	0.17	0.20	**1.00**	2.50	3.00	3.50	3.60	2.83	0.25	0.50	0.80
Man-HLL	USA	28	0	0.88	0.75	0.80	2.25	**1.50**	3.00	3.17	3.20	1.50	0.75	1.00	1.00
Man-HLL	Canada	19	0	1.00	0.92	0.60	**1.50**	1.75	2.40	**1.50**	2.00	1.50	0.75	0.50	1.00
Can-HLL	China	23	4	1.00	1.00	0.50	2.00	2.25	**1.80**	**1.33**	1.80	1.50	1.00	0.75	1.00
Multivariate outliers[h]															
Can-HLL	Hong Kong	20	0	1.00	0.75	1.00	3.25	3.25	4.00	4.00	2.20	3.00	0.50	0.75	1.00

- AoA= age of arrival in an English-speaking country
- Comp.= comprehending strategies
- Monitor.= monitoring strategies
- Eval.= evaluating strategies
- OrthoP= ortho-phonological task
- OrthoS= ortho-semantic task
- The threshold of univariate outliers is three standard deviations away from the mean.
- The threshold of multivariate outliers is $p<0.001$ using Mahalanobis distance (D).

All the test takers with missing data were deleted from the dataset. Pairwise scatterplots, shown in Figure 5, were checked for nonlinearity and heteroscedasticity. In Figure 5, the pairwise scatterplots were shown on the upper side with the corresponding Pearson correlation coefficients on the lower side. Nonlinearity and low correlations were identified in some scatterplots. According to Tabachnick and Fidell (2012), a sample size as small as the one used in the present study could lead to nonlinearity and endanger the power of the comparisons. Therefore, when interpreting the results of comparisons, it should be borne in mind that group differences are likely to be more difficult to detect, due to lower power.

Another potential risk examined through pairwise scatterplots was heteroscedasticity, which is related to Singaporean CMTLLs' better performance on the reading test, in that discrepant distributions were likely to inflate the relationship between any pair of subtests. However, no discrepant distributions could be identified in the scatter plots shown in Figure 5, indicating that Singaporean CMTLLs' test performances were less likely to inflate the following statistical analyses.

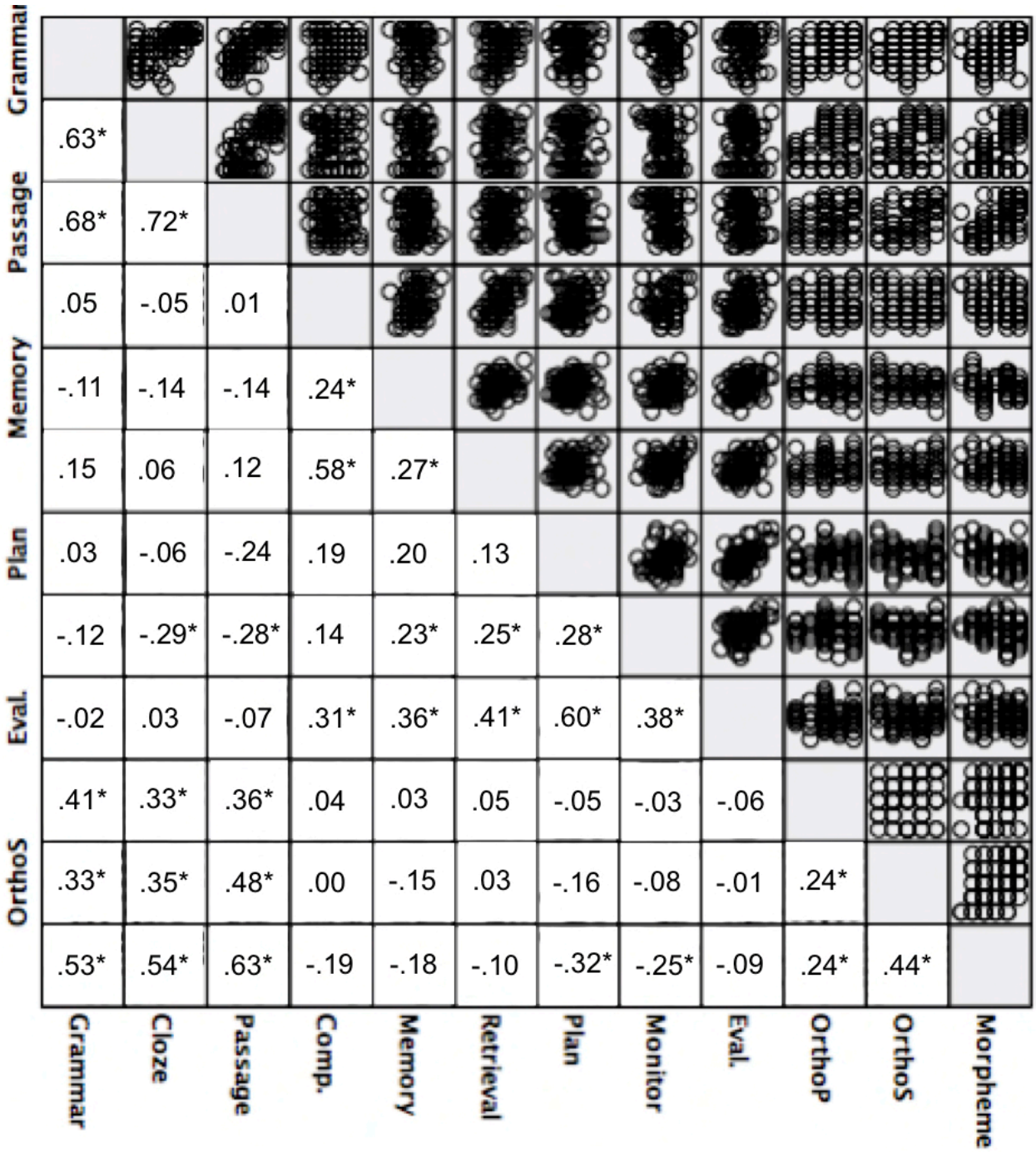


Figure 5. Bi-variate scatter plots for the 12 variables. Comp. = comprehending strategies, Plan. = planning strategies, Monitor. = monitoring strategies, Eval. = evaluating strategies, OrthoP = ortho-phonological subtest, OrthoS = ortho-semantic subtest. * = $p < 0.05$

Table 15

Descriptive Statistics of Each Subtest

n = 85 Subtest	*k*	Possible Range	Min	Max	*Mean*	*SD*	Skewness SE = 0.26[d]	Kurtosis SE = 0.52[d]
Grammar	8	0-1	0.00	1.00	0.68	0.26	-0.63	-0.64
Cloze	12	0-1	0.00	1.00	0.45	0.34	-0.02	-1.31
Passage	10	0-1	0.10	1.00	0.58	0.28	0.01	-1.32
Comp.[a]	4	1-4	2.25	4.00	3.10	0.46	0.37	-0.61
Memory	4	1-4	1.75	4.00	2.87	0.40	-0.08	0.24
Retrieval	5	1-4	2.40	4.00	3.11	0.35	0.59	0.08
Planning	6	1-4	2.00	3.83	2.84	0.37	0.36	0.51
Monitoring	5	1-4	1.80	4.00	3.01	0.42	-0.29	0.51
Evaluating	6	1-4	1.83	4.00	2.82	0.44	0.30	0.55
OrthoP[b]	4	0-1	0.00	1.00	0.58	0.32	-0.33	-0.74
OrthoS[c]	4	0-1	0.00	1.00	0.56	0.33	-0.20	-1.06
Morpheme	5	0-1	0.00	1.00	0.65	0.29	-0.44	-0.99

a. Comp.= comprehending strategies
b. OrthoP= ortho-phonological task
c. OrthoS= ortho-semantic task
d. The alpha levels were set at 0.001 for small samples, as recommended by Tabachnick and Fidell (2012)

The descriptive statistics, skewness, and kurtosis of the 12 subtests are presented in Table 15. Detailed descriptive statistics of all the items are included in Appendix B. All the skewness and kurtosis figures were within the range of mean ± 3.31 standard deviations (Tabachnick & Fidell, 2012), so no transformation of the scores was necessary. Multivariate outliers were checked using Mahalanobis distance, and a multivariate outlier (see Table 14) was dropped from the analyses. Multicollinearity and singularity were checked and all the correlations were smaller than 0.90, which met the suggested threshold (Tabachnick & Fidell, 2012). At the end of data screening, 85 participants remained in the dataset. Then, the reliability of all the variables was computed using Cronbach's alpha, and the results are shown in Table 16. Reliabilities of the six strategy-use are lower than ideal and the lower reliabilities may be results of few items in each of

the strategy categories and the standard deviations of the six strategy categories were small as well (see Table 11). Comparing the reliability coefficients between the second pilot session and the main study, most of the reliability coefficients improved from the second pilot session, but the reliability of comprehending and planning strategies decreased in the main study. Furthermore, although the reliability of memory strategy improved, its internal consistency was still in question. Considering that most of low reliability coefficients were from strategy-use survey, it is likely that the non-controlled testing contexts and unclear reading purposes lead to the low reliability in these strategy types.

Table 16

Summary of Reliability Coefficients

	Cronbach's Alpha	
	k	Reliability
Grammar	8	0.72
Cloze	12	0.91
Passage	10	0.79
Comp.[a]	4	0.65
Memory	4	0.31
Retrieval	5	0.44
Planning	6	0.52
Monitoring	5	0.60
Evaluating	6	0.75
OrthoP[b]	4	0.69
OrthoS[c]	4	0.61
Morpheme	5	0.60

a. Comp.= comprehending strategies
b. OrthoP= ortho-phonological task
c. OrthoS= ortho-semantic task

3.6.3 Profile analysis. Profile analysis is a special type of multivariate analysis of variance (MANOVA), which examines the influence of grouping variable(s) on multiple outcome variables (Tabachnick & Fidell, 2012). The advantage of MANOVA over univariate analysis of variance (ANOVA) is that MANOVA automatically adjusts for Type-I error through

a Bonferroni approach, and thereby generates more robust comparison results than ANOVA does. The main advantage of profile analysis over MANOVA is its graphic presentation of the comparisons across variables (Tabachnick & Fidell, 2012). In the current study, profile analysis is used for examining the four levels of group membership over multiple independent variables (IVs) at the same time. To be more specific, the four levels of group membership (Singaporean CMTLLs, CFLLs, Man-HLLs, and Can-HLLs) were treated as dependent variables (DVs), while the IVs were the participants' scores on the 12 subtests. Three separate profile analyses were performed, and the experimentwise Type-I error rate was set at 0.05. In these three analyses, the 12 IVs were grouped based on three types of reading processes: i.e., the first type of reading skills included the three subtests of the reading test, representing reading comprehension ability; the second type included the three subtests of the character-recognition skills test, representing character-recognition ability; and the third type included the six strategy categories, representing strategy use as an aspect of higher-level reading processes.

Prior to running profile analyses, data should be screened to ensure that certain assumptions are all met. These assumptions are multivariate normality, absence of univariate and multivariate outliers, homogeneity of variance-covariance matrices (confirmed by nonsignificant Box' M test results), linearity, multicollinearity, and singularity. In the present study, all these assumptions with the exception of homogeneity were checked in the previous phase of data screening (see Table 13). The assumption of homogeneity was violated in one of the three profile analyses, and the follow-up steps will be described in the next chapter.

To determine the sources of significant differences in the results, profile analysis conducts three types of comparisons: profile flatness, profile parallelism, and overall difference across groups, and these comparisons are conducted with the flatness test, the parallelism test,

and the level test, respectively. The flatness test examines the differences among the mean scores of the DVs without regard to grouping variables; the parallelism test inspects interactions between group membership and subtests; and the level test examines differences among the group means of the total score, without regard to subtests (Tabachnick & Fidell, 2012). The results of the level test and the parallelism test were the primary foci in the efforts to answer the first research question.

In addition to significance values, effect sizes were included as a measure of how meaningful group differences were. To reflect the effect sizes of the group differences, partial η^2 (eta-squared) was adopted. Partial η^2 is similar to R^2 in multiple regression, in that both are used to explain the overlapping variance between IVs and DVs; and partial η^2, rather than η^2, is adopted in profile analysis because the IVs in this study are not independently measured (Brown, 2008). For example, if partial η^2 is 0.50, it means that 50% of the variance in the DVs can be explained by the IV(s).

3.6.4 Structural Equation Modeling (SEM). SEM is a combination of multiple regression, factor analysis, and path analysis (Tabachnick & Fidell, 2012). Similar to multiple regression, SEM is a correlational analysis examining covariance between variables; it includes path diagrams depicting causal relationships between variables and factor loadings of the paths. However, because SEM is a correlational analysis, its path diagrams are required to be theory-driven. In other words, SEM is best used for *a priori* studies (Mueller & Hancock, 2010). Mueller and Hancock also suggest that SEM should be treated as an analytical process rather than as a mere statistical method, and as such, it should be broken down into model conceptualization, parameter identification and estimation, data-model fit assessment, and

potential model re-specification. Each of these phases, as it applies to the present research, is described more fully below.

3.6.4.1 Model conceptualization. Because SEM mainly examines a hypothesized model based on correlation, the causal relationships of path diagrams should be supported by theories and other experimental or empirical studies examining such theories (Byrne, 2010; Mueller & Hancock, 2010). In the graphic representation of an SEM model, a square represents an observed variable or a factor, and a circle is something unobserved in the study, referred to as a latent variable. Arrows between variables or latent variables denote their causal relationships: with the variable being pointed at being an effect of another variable, and the variable where the arrow originates, a cause. In the graphic representation of an SEM model, the part comprising measurement variables and their related latent variable is defined as a measurement model, and the part consisting of the paths among latent variables is defined as a structural model. When estimating the correlations in a structural model, error variance of the measurement variables is excluded. In other words, only the common and unique variances enter a structural-model estimation (Mueller & Hancock, 2010; Tabachnick & Fidell, 2012).

Taking the hypothesized model in the current study as an example, variance was observed in 12 subtests, and these subtests were shown in squares in Figure 1 in Chapter 2. On the other hand, reading comprehension ability, character-recognition ability, use of cognitive strategies, and use of metacognitive strategies were not directly observed, and were therefore conceptualized as latent variables, and shown in circles. The causal relationship between the variables can be explained by the latent variable of reading ability, where the arrows originate; these arrows point at the three subtests (the grammar subtest, the cloze subtest, and the passage-

comprehension subtest), indicating that the variances in the three subtests were caused by the variance in participants' reading comprehension ability.

Other than variance resulting from the latent variable, variance in the observed variables may also be caused by measurement errors, and the measurement error variances are also accounted for in the hypothesized model. In addition to measurement errors, there may be *residual errors*, i.e., unexplained variances that arise when predicting the variance of latent variables with the variances of observed variables (Byrne, 2010; Kline, 2011). Mueller and Hancock (2010) note that, ideally, a latent variable should be estimated based on more than three observed variables, as this will lead to higher stability across samples; and if a latent variable has too many indicators – generally, more than six – researchers can compile these indicators into a single composite score.

3.6.4.2 Parameter identification and estimation. After conceptualizing the SEM model, it is important to set its degree of freedom at a level equal to or higher than zero, and to place all of its latent variables, including their residual loadings, on the same scale (Kline, 2011). The degree of freedom is higher than zero if the number of estimated parameters is smaller than $p(p+1)/2$, where p equals the number of observed variables; and the estimated parameters are those in the hypothesized SEM model that are freely estimated, or whose values are not set at one (Kline, 2011; Ockey, 2014). To assign all latent variables to the same scale, each measurement model needs a parameter to be set at one. It is recommended that the parameter connecting the latent variable and the observed variable with highest reliability in a measurement model to be set at one, and keep the remaining parameters in the same measurement model freely estimated (Ockey, 2014). To estimate the loading of each error and residual, the most common practice is to set the loading at one (Kline, 2011). Sometimes, however, an error variance is

estimated as negative, and in such cases, researchers need to set the error variance at zero or slightly above zero to avoid this (Mueller & Hancock, 2010).

Maximum likelihood (ML) is the most widely recommended estimation method. However, researchers should also be aware of its limitations: especially that its estimates of factor loadings and model-fit index (χ^2, chi-square), are inflated in the case of a sample with non-normal distribution (Mueller & Hancock, 2010). Non-normal distribution can be detected by skewness (> 2 or < -2), kurtosis (> 7 or < -7), and Mardia's normalized coefficient of multivariate kurtosis (> 1.96). Multivariate outliers can also be examined using Mahalanobis distance ($p < 0.001$) (Mueller & Hancock, 2010).

After the parameters of an SEM model are identified, SEM provides unstandardized and standardized estimates of each of them, along with standard errors and *p*-values of the unstandardized estimates. Ideally, parameter estimates should be statistically significant, but nonsignificant parameter estimates may simply be a function of small sample size (Byrne, 2010).

3.6.4.3 Data-model fit assessment. It is important to examine model-fit indices if one is to understand how well the hypothesized model fits the data. A popular quick check of model fit is χ^2, which examines the difference between the hypothesized model and the data; in other words, the higher the value of χ^2, the worse the hypothesized model fits the data. When a hypothesized model adequately fits the data, the *p*-value of the χ^2 value with its degree of freedom should be statistically nonsignificant ($p > 0.05$) (Byrne, 2010). However, because χ^2 statistics are sensitive to sample size, other model-fit indices should also be included when reporting model fit. Mueller and Hancock (2010) divided model-fit indices into three categories: absolute, parsimonious, and incremental. Standardized root mean square residual (SRMR) is regarded as an absolute index, and represents the discrepancy between the implied covariance in

the hypothesized model and the observed covariance in the data; ideally, SRMR should be lower than 0.08. The root mean square error of approximation (RMSEA) is a parsimonious index; it is similar to SRMR, but takes model complexity into account. RMSEA and its range of 90% confidence interval (higher bound minus lower bound) should be below 0.05, and the lower bound of the 90% confidence interval should be equal to zero (Kline, 2011; Mueller & Hancock, 2012). Furthermore, even if an RMSEA is lower than 0.05, its estimate may not be ideally precise if it has a 90% confidence interval wider than 0.05.

Comparative fit index (CFI) and Tucker-Lewis index (TLI) are both incremental indices. They reflect how much better the hypothesized model fits the data than a null model, or the independent model, does. Both are adjusted for the effect of the sample size. The null model assumes no relationship between the variables in the model, and thus the higher CFI and TLI are, the better the hypothesized model is (relative to the null model) in terms of its fit to the data. The key difference between CFI and TLI is that CFI is a normed fit index whereas TLI is non-normed, but the value of both CFI and TLI should be higher than 0.95 (Byrne, 2010; Mueller & Hancock, 2010).

3.6.4.4 Potential model re-specification. After examining how well the hypothesized model fits the data, ML provides modification indices suggesting how model fit could be improved by adding more paths in the model. However, it must be noted that this kind of model re-specification is exploratory, and additional theoretical consideration is necessary when adding or dropping parameters in the model (Mueller & Hancock, 2010). At this exploratory step, according to Byrne (2010), researchers should try to identify the source of misfit, and also consider alternative SEM models. The same author proposed three considerations when re-specifying a hypothesized model. First, the inclusion of additional parameters must be

theoretically supported. Second, the inclusion of additional parameters may lead to the modified model overfitting, i.e., becoming too specific to the data to reflect the constructs of the larger population. And third, if the value of an expected parameter change for the modification index is substantial, this should be taken as an indication that including the parameter would improve the overall quality of the fixed parameters in the model.

The risk of overfitting is that the statistical evidence emerging from the data may be a trend that is only be observable in the current data, and might not be replicable across the whole target population (Mueller & Hancock, 2010). Similarly, although some SEM computer programs provide modification indices that suggest the dropping of nonsignificant parameters, Mueller and Hancock have recommended keeping such parameters, as they are theory-driven at the model-design stage, and acknowledging that the relationships in the target population/context may exist simply by chance. Lastly, after re-specifying a hypothesized model, researchers should discuss both the statistical and theoretical justifications for the replicability of the modified model in the wider target population (Mueller & Hancock, 2010).

After an SEM model is re-specified, it is also necessary to investigate how much it has improved upon the initial one. Mueller and Hancock (2010) have suggested that when the initial model is nested in the revised model, χ^2 difference (with degree-of-freedom difference) should be examined, with significant results indicating that the two models are statistically different. For example, the χ^2 difference test for an initial model (χ^2 = 89, df = 10) and a revised model (χ^2 = 80, df = 9) yields a significance level for χ^2 = 89 - 80 = 9, df = 10 - 9 = 1, which is $p < 0.005$, indicating that the revised model is significantly better than the initial model.

3.6.5 Exploratory factor analysis. Like SEM, exploratory factor analysis (EFA) examines correlations between variables, analyzes which variables show strong correlations with

each other, and then extrapolates from these observed variables to a smaller number of unobserved factors. Usually, EFA is conducted before SEM to confirm the extent of the correlations within the data. However, in this study, EFA was conducted after analyzing the hypothesized SEM model, for the purpose of SEM model re-specification. Moreover, the rationale for choosing EFA over other methods of factor analysis, such as principal component analysis (PCA), is that the literature strongly supports the inclusion of the observable variables and their relationships to the corresponding latent variables in the present study.

EFA includes five main steps, which are: screening and preparing the data, extracting the factors, determining the number of extracted factors, rotating the factors to increase interpretability, and interpreting the resultant factors (Ockey, 2014; Tabachnick & Fidell, 2012); and a more detailed explanation of each of these five steps is set forth below. When screening and preparing the data, assumptions of normality, linearity, multicollinearity, and singularity should all be met, and there should be no univariate or multivariate outliers. In addition, to ensure the factorability of the correlation matrix, Kaiser's measure of sampling adequacy and the anti-image matrix should both be examined. The value of Kaiser's measure of sampling adequacy should be higher than 0.6, and most of the off-diagonal elements of the anti-image matrix should be small (Tabachnick & Fidell, 2012).

After screening and preparing the data with Kaiser's measure of sampling adequacy and the anti-image matrix, the most common extraction method for EFA is principal axis factoring, which analyzes only the common and unique variance, and excludes error variance (Brown, 2009; Ockey, 2014; Tabachnick & Fidell, 2012). Another common extraction method is PCA, in which the overarching constructs are defined as *components* rather than *factors*. The difference between a factor and a component is that the former is extracted from a correlation matrix in

which the diagonals denote only covariance without error variance, whereas a component is extracted from a correlation matrix containing the value of one in the diagonals of the correlation matrixes, indicating all the variance is analyzed (Kline, 1994; Tabachnick & Fidell, 2012). In other words, a factor is extracted when error variance is excluded from among a set of variables; thus, it is more useful for testing theories that can be applied to a wide range of contexts. On the other hand, an extracted component may be viewed as contaminated by error variances, which limits its application to a broader context (Kline, 1994). Another difference between a factor and a component is that EFA should be theory-driven, whereas the theoretical support for PCA may be of a lesser degree (Tabachnick & Fidell, 2012). When EFA extracts factors, it first forms a linear combination of variables to maximally explain the variance. This linear combination of variables is the first factor, and the weight of each variable in it is the *factor loading* (see next paragraph) of the variable on the component. After the first factor is extracted, other factors are extracted iteratively (Kline, 1994; Tabachnick & Fidell, 2012).

Factor loadings are the correlations between the observed variables and the unobserved factors, and should always be included when reporting the extracted factors. The average of the squared factor loadings should also be reported, as it indicates the percentage of variance explained by the factor; and so should the sum of the squared factor loadings in a row, defined as commonalities or h^2 (Brown, 2009; Kline, 1994).

It is important to keep the number of extracted factors as low as possible and still reliably explain total variance. The first criterion to consider when deciding on the number of extracted factors is that the selected ones should all have eigenvalues larger than one (Tabachnick & Fidell, 2012). A factor's eigenvalue represents the extent of the variance explained by it, so if an eigenvalue is less than one, the associated factor is not representative as an overarching construct

of the variables (Kline, 1994). The second criterion is based on the scree plot of the eigenvalues. Because the first factor gets the lion's share of the variance, the eigenvalue of the first factor is the highest, and the eigenvalues of the successive factor keep decreasing, leading a scree plot of any set of eigenvalues to have a downward slope. However, when the eigenvalue slope falls away suddenly, this indicates that the variance explained by the successor of the factor in question is substantially less than that by the component in question. In other words, only the factors before this slope-angle turning point are extracted, and the factor at which the turning point occurs should be extracted (Tabachnick & Fidell, 2012). However, it is often the case that there are multiple turning points, or none. Gorsuch (1974) explained that these indeterminate scree plots can result from small samples and non-random data. However, given a large enough random sample, the number of extracted factors arrived at by examining scree plots is usually suitable to the theory on which the measurements have been developed.

Because EFA extracts maximum variance for the first factor, many variables tend to load highly on the first factor (especially in cases where the explained variance is large), and this leads to difficulties in interpretation. Therefore, to increase factor interpretability, rotation is necessary. Rotation is divided into two types: orthogonal, and oblique (Kline, 1994; Tabachnick & Fidell, 2012). Orthogonal rotations treat factors as uncorrelated; however, in the social sciences, factors (and components) are seldom uncorrelated. As such, oblique rotations are more useful in social-science contexts because they allow for factors to be correlated with each other to varying degrees (Kline, 1994).

Nonetheless, the use of oblique rotations also leads to certain difficulties in interpreting, describing, and reporting rotated factors (Tabachnick & Fidell, 2012). First, the sum of the squared factor loadings in each row could be equal to the factor's commonality simply due to

chance. Second, the average of the squared factor loadings may not be equal to the percentage of variance explained by the factor. Third, the sum of the oblique loadings could be equal to the total variance in the matrix by chance (Kline, 1994). Varimax is the most commonly used type of orthogonal rotation among researchers interested in simpler factor structures, but for those who assume that their factors are correlated, the recommended method of oblique rotation is direct oblimin with delta value equal to zero, allowing a moderate degree of correlations among extracted factors (Field, 2013; Tabachnick & Fidell, 2012).

When reporting the results of oblique rotations, the pattern matrix and the factor (or component) correlation matrix should be included (Tabachnick & Fidell, 2012). In a pattern matrix, the factor loadings represent the unique contributions of the variable to the factor, and are analogous to beta-weights in multiple regression. These factor loadings remove the overlapping variance between factors, leading to clarification of the relationships between variables and factors (Tabachnick & Fidell, 2012). When interpreting factors based on a pattern matrix, only factor loadings over 0.32 should be considered. Each variable should have only one high factor loading on a factor, and the remaining factor loadings should be close to zero; and the number of variables showing high loadings on more than one component should be minimal (Kline, 1994). Lastly, a content analysis of the variables is necessary to describe the extracted and rotated factors (Ockey, 2014).

This chapter has described the present study's measurement procedures, measurement instruments, participants' backgrounds and subtest scores, and analysis methods. In summary, the measurement instruments designed for the current study included three reading subtests, a strategy-use survey consisting of six strategy types, and three character-recognition subtests, representing the 12 study variables. In total, 85 participants remained in the participant pool, and

were divided into four groups, rather than the three called for in the original study design. In general, among these four groups, Singaporean CMTLLs had received the longest formal education in Mandarin and exhibited the highest mean scores on the three reading subtests and the three character-recognition subtests. Man-HLLs, meanwhile, reported the most Mandarin use at home, and scored the second highest on the reading and character-recognition subtests. There were no clear group differences in strategy-use survey responses between CFLLs and Can-HLLs. Three profile analyses were conducted to answer the first research question, and SEM and EFA were adopted to answer the second and third research questions.

The next chapter first summarizes the descriptive statistics of the participants' test scores, and then reports the results of profile analysis and SEM by way of answering the three research questions.

CHAPTER 4

RESULTS

The present study aims to attain a better understanding of three aspects of L2 Chinese reading: language background manifested as a source of background knowledge; character-recognition skills, as a lower level of reading-processing skills; and strategy-use and reading comprehension skills, as two higher levels of those skills.

4.1 First Research Question: To What Extent Can Language Background Explain the Variance in Performance on the Reading Comprehension Test, the Character-Recognition Test, and the Strategy-Use Survey?

As previously discussed, the participants were divided into four groups, despite the original study design having included only three groups, because the descriptive statistics suggested that Singaporean CMTLLs were sufficiently different from the other three groups that this was advisable. To answer the first research question, three separate profile analyses were conducted, on the reading test, the character-recognition test, and the strategy-use survey. To adjust for the inflated Type-I error caused by the three profile analyses, the critical *p*-value of each profile analysis was divided by three (0.05/3 = 0.017). The software used to conduct the profile analyses was IBM SPSS Statistics 21.

4.1.1 First profile analysis. Language background was the DV in all the profile analyses, and the four levels of this DV were Singaporean CMTLLs, CFLLs, Man-HLLs, and Can-HLLs. In the profile analysis that examined the reading test, the three IVs were the grammar subtest, the cloze subtest, and the passage-comprehension subtest. All the assumptions of profile analyses were checked as part of the procedure of data screening. Violation of the homoscedasticity assumption was found, which could be caused by unequal sample sizes, the homoscedasticity

should be assessed with an F_{max} test when a study has unequal sample sizes. The F_{max} value should be less than is 10 (Tabachnick & Fidell, 2012, p. 86). For the grammar subtest, the F_{max} was 15.60; for the cloze subtest, it was 4.63; and for the passage-comprehension subtest, 1.63. If the homoscedasticity assumption is violated, Tabachnick and Fidell (2012) have recommended either transforming the raw scores, or setting a lower critical *p*-value (e.g., adjusting 0.05 to 0.025 or 0.01). If a researcher chooses to transform the raw scores based on the same authors' advice, then negatively skewed data – like the distribution of the grammar subtest – should be reflected and square-rooted. In addition, the result interpretations should be limited to transformed data, making it more of a challenge to generalize the findings to the target population. Hence, I chose to the second option, setting the Type-I error rate at 0.0085 (0.017/2), to adjust for the violation of the homoscedasticity assumption in the grammar subtest.

In addition, the test of sphericity, $p = 0.015$, indicated a violation of the homogeneity of covariance,[7] mandating that the level test and the parallelism test be adjusted using Greenhouse-Geisser (G-G) test or Huynh-Feldt (H-F) test (Tabachnick & Fidell, 2012). Whether G-G or H-F testing is chosen should be based on the G-G estimate provided by the test of sphericity. If this G-G estimate is 0.75 and lower than to 0.75, then the G-G test should be used, but if it is higher than 0.75, then the H-F test should be utilized instead (Field, 2013). In this profile analysis, because the G-G estimate in the test of sphericity was 0.91, the results were adjusted by an H-F test, as reported in Table 17.

[7] The homogeneity of covariance assumption holds that all possible comparisons of the level test and the parallelism test are equally correlated (Tabachnick & Fidell, 2012).

Table 17

Profile Analysis for Reading Subtests by Groups

Critical *p*-value: 0.0085	*SS*	*df*	*MS*	*F*	*p*	Partial η^2	Observed Power
Between Groups							
Groups (levels)	3.289	3	1.096	6.027	0.001	0.182	0.951
Error	14.735	81	0.182				
Within Groups							
Subtests (flatness)	2.624	1.926	1.362	43.474	< 0.001	0.349	1.000
Subtests * Group (parallelism)	0.451	5.778	0.078	2.488	0.027	0.084	0.813
Error	4.890	156.016	0.031				

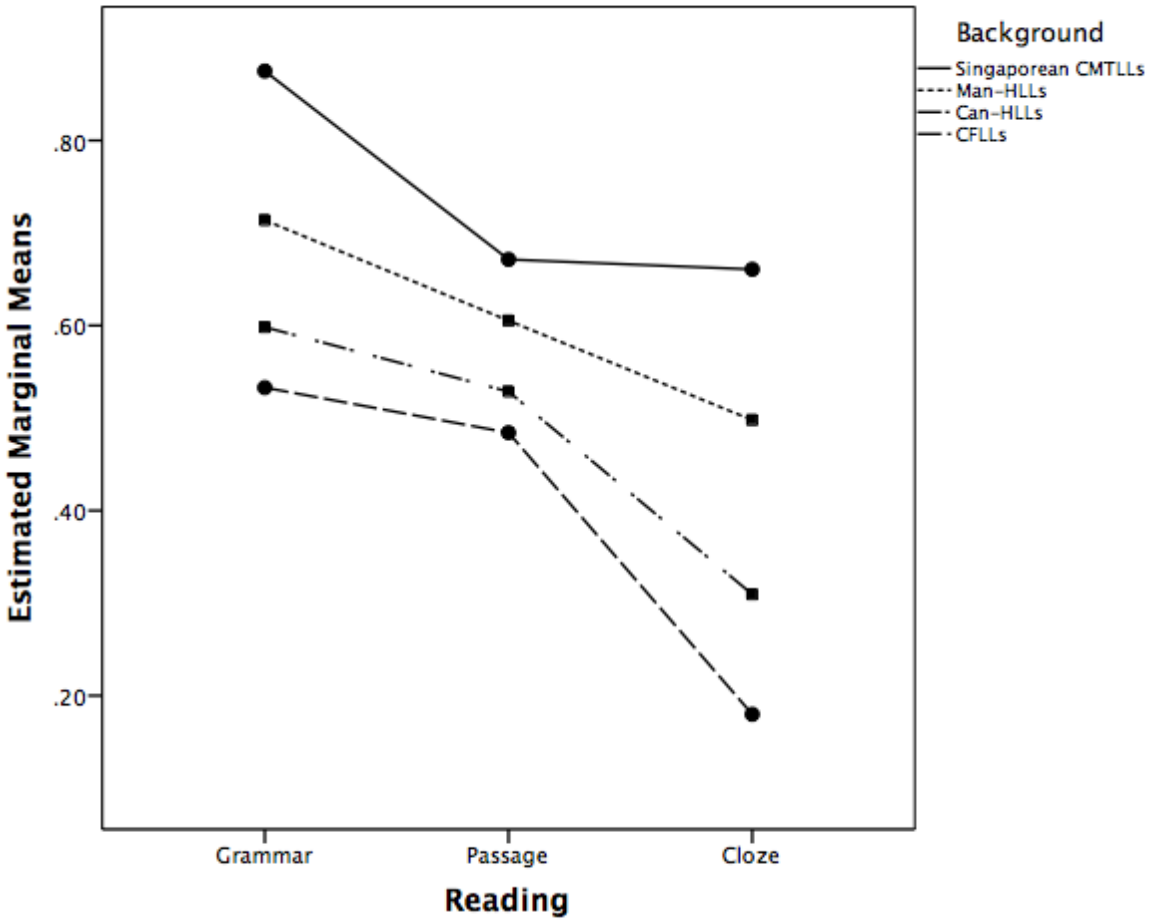


Figure 6. Profiles of reading subtests for the four groups

From the results of the flatness test, it can be seen that the profile of each subtest performed differently from each of the others, and that such differences were significant: $F(3, 5.778) = 43.474$, $p < 0.001$, partial $\eta^2 = 0.349$. However, the profiles did not show significant deviation from parallelism, $F(3, 5.778) = 2.488$, $p = 0.027$, partial $\eta^2 = 0.084$. The level test found significant differences among the four groups, $F(3, 81) = 6.027$, $p = 0.001$, partial $\eta^2 = 0.182$.

Figure 6 provides a more intuitive picture of the group differences across the three subtests, arranging based on subtest difficulties. Each of the four lines represents a group's profile. First, the line representing the Singaporean CMTLLs exhibits the highest scores on the three subtests, with Man-HLLs second, Can-HLLs third, and CFLLs the lowest. Second, the four lines are almost parallel to each other, meaning that no significant interactions between group membership and subtest difficulty were found; in other words, for all four groups, the grammar subtest was the easiest, the passage-comprehension was the second easiest, and the cloze subtest was the most difficult. Third, the cloze subtest appears to have elicited the widest inter-group differences of any of the three subtests, while the passage-comprehension subtest elicited the narrowest differences.

In sum, the variance in test performance caused by language background is indicated by the effect size of the level test, i.e., approximately 18% of the total variance in reading-test scores. The greatest amount of the total variance, around 35%, was explained by differences among the subtest difficulty levels; and no significant interaction between group membership and subtest was found.

4.1.2 Second profile analysis. As in the first profile analysis, language background was the DV, while the three IVs were the ortho-phonological subtest, the ortho-semantic subtest, and

the morpheme subtest. All the assumptions of profile analyses were met. In addition, the result of the test of sphericity, $p = 0.051$, indicates that the assumption of homogeneity of covariance was met.

Table 18

Profile Analysis for Character-Recognition Subtests by Groups

Critical *p*-value: 0.017	*SS*	*df*	*MS*	*F*	*p*	Partial η^2	Observed Power
Between Groups							
Groups (levels)	3.289	3	0.401	2.689	0.052	0.091	0.521
Error	14.735	81	0.149				
Within Groups							
Subtests (flatness)	0.282	2	0.141	2.113	0.064	0.066	0.427
Subtests * Group (parallelism)	0.718	6	0.120	1.790	0.058	0.072	0.637
Error	10.827	162	0.067				

Based on the flatness test, as shown in Table 18, it can be seen that the profiles of these three subtests were not significantly different from each other, $F(3, 2) = 2.113$, $p = 0.064$, partial $\eta^2 = 0.066$ (using Wilks' criterion). Also, the profiles did not show significant deviation from parallelism, $F(3, 6) = 1.790$, $p = 0.058$, partial $\eta^2 = 0.072$ (using Wilks' criterion). The level test found significant differences among the four participant groups, $F(3, 81) = 2.689$, $p = 0.052$, partial $\eta^2 = 0.091$.

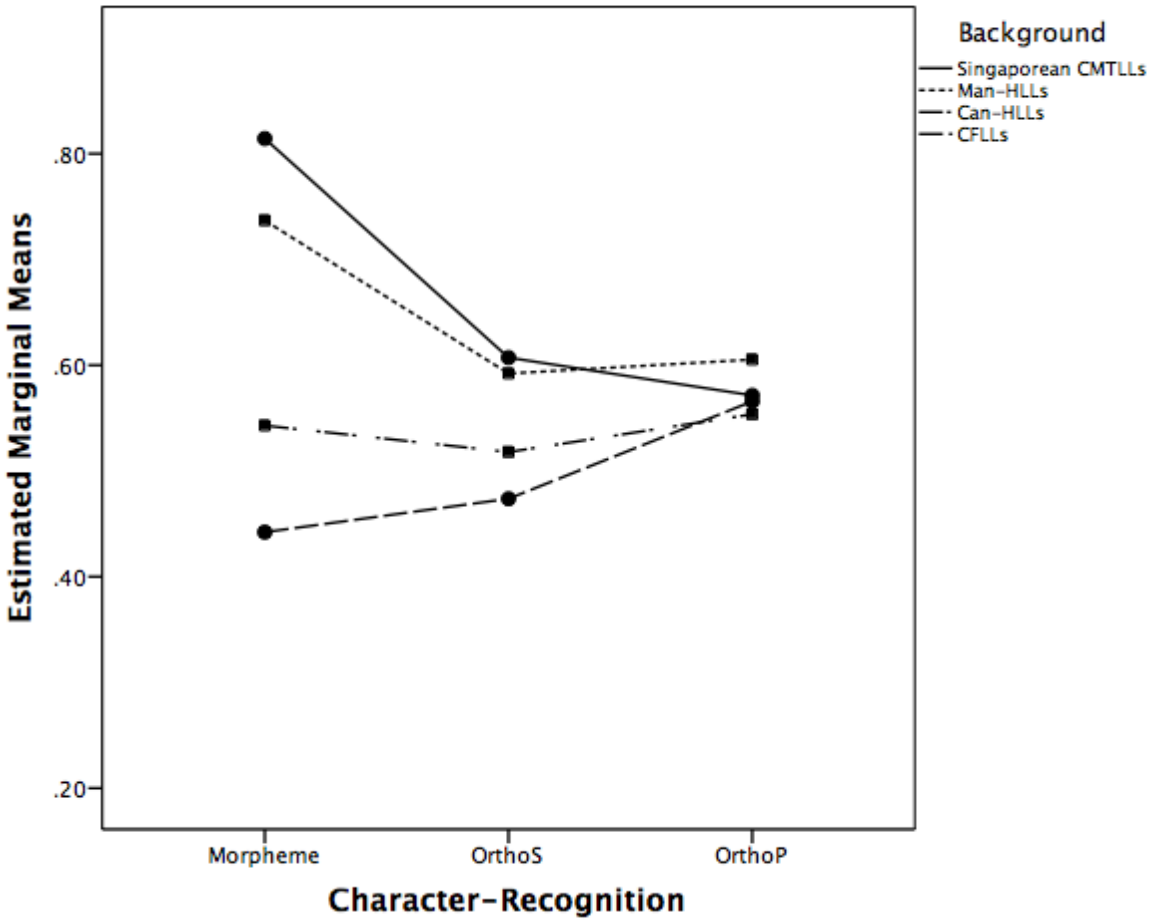


Figure 7. Profiles of character-recognition subtests for the four groups

Additionally, Figure 7 shows that Singaporean CMTLLs exhibited the highest scores, with Man-HLLs second, Can-HLLs third, and CFLLs the lowest. The morpheme subtest seems to have revealed the largest group difference, and the ortho-semantic subtest the second largest, while almost no group differences were discerned by the ortho-phonological subtest. In terms of character-recognition subtest performance, Singaporean CMTLLs and Man-HLLs were similar to each other, and Can-HLLs and CFLLs were also similar to each other. In this case, the four lines are not parallel to each other, and their divergence is especially marked between the character- and the subcharacter levels, i.e. ortho-semantic and ortho-phonological skills. The non-parallel lines indicate interactions between group membership and character-recognition

skills; and more specifically, that the effect of group membership may be stronger at the character level than at the subcharacter level (this being the cause of the crossed lines in Figure 7).

In sum, no significant differences were found in the second profile analysis. On the other hand, although no significant differences were found among the three subtests, the effect size of the level test suggests that approximately 10% of total variance in the character-recognition test scores can be explained by group differences. It is noteworthy that the level test found the largest difference among the three comparisons, even though its effect size is small.

4.1.3 Third profile analysis. In the third profile analysis, language background was the DV, and the six IVs were comprehending, memory, retrieval, planning, monitoring, and evaluating strategies. All the assumptions of profile analyses were checked as part of the data-screening procedure. In addition, the test of sphericity ($p = 0.001$) and G-G estimate (0.86) indicate that the assumption of homogeneity of covariance was violated, and that an H-F test was needed to adjust for the inflated Type-I error.

Table 19

Profile Analysis for Strategy-Use by Groups

Critical *p*-value: 0.017	*SS*	*df*	*MS*	*F*	*p*	Partial η^2	Observed Power
Between Groups							
Groups (levels)	1.700	3	0.567	1.372	0.257	0.048	0.251
Error	33.447	81	0.413				
Within Groups							
Subtests (flatness)	7.013	4.730	1.483	12.325	< 0.001	0.132	1.000
Subtests * Group (parallelism)	2.694	14.191	0.190	1.578	0.081	0.055	0.807
Error	46.09	383.144	0.120				

From the flatness test, as shown in Table 19, the profile of each strategy type performed significantly differently from each of the others: $F(3, 4.730) = 12.325$, $p < 0.001$, partial $\eta^2 = 0.132$. However, the profiles did not show significant deviation from parallelism, $F(3, 14.191) = 1.278$, $p = 0.081$, partial $\eta^2 = 0.055$. Also, the level test did not find significant differences among the four groups, $F(3, 81) = 1.372$, $p = 0.257$, partial $\eta^2 = 0.048$. The effect size found by the level test indicates that approximately 5% of total variance in the participants' strategy-use was related to group differences.

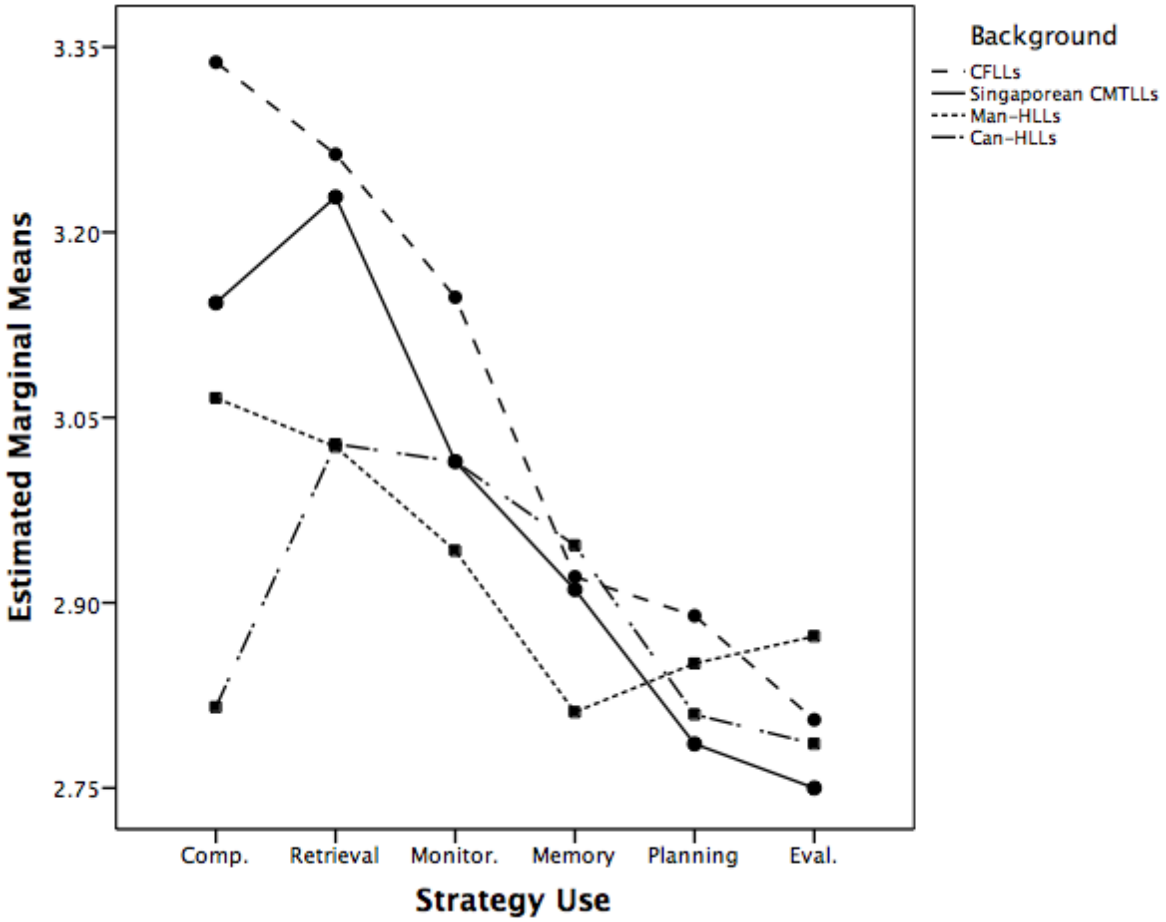


Figure 8. Profiles of the strategy-use survey for the four groups. Comp. = comprehending strategies, Monitor. = monitoring strategies, Eval. = evaluating strategies.

Figure 8 provides a graphic representation of group differences across the six strategy types, arranged based on strategy use. First, it can be seen that the four lines are not horizontal, indicating the statistical differences found in the flatness test and explaining the source of the largest variance (13%) in the third profile analysis. Second, the line representing CFLLs appears to indicate that this group had the most positive attitude to their strategy use while taking the reading test. Third, cognitive strategies, especially comprehending strategies, appear to demonstrate larger group differences than metacognitive strategies, which are planning, monitoring, and evaluating strategies, do. And lastly, although the four lines seem to cross each other several times in Figure 8, no significant interaction between group membership and strategy use was found by the third profile analysis. This lack of statistical significance may be because most of the participants' attitudes toward their strategy-use scored between 2.75 and 3.35. This is a narrow range on a 4-point scale, and the interactions as indicated by line-crossing are thus less likely to be statistically significant.

To conclude the three profile analyses, after adjusting the *p*-values to appropriate levels for the multiple comparisons made, significant group differences were only found in regard to the reading test, where such differences explained about 18% of the total variance in scores. In regard to responses on the character-recognition test and the strategy-use survey, no significant group differences were found, and such differences as did exist explained less than 10% of the variance in the former and less than 5% of the variance in the latter. The parallelism test found significant interaction between group membership and subtest only in the second of the three profile analyses, implying that the significant interaction between group membership and character-recognition skills shown in Figure 7 explains approximately 7% of the total variance. The flatness tests of the reading test and the strategy-use found the largest difference among in

the three profile analysis, suggesting that subtest difficulty and strategy type is a more important source than language background in terms of explained variance. Lastly, to answer the first research question, across the three profile analyses, group difference explains around 5% to 18% of total variance, and interaction between group membership and subtest type explains around 6% to 8% of total variance.

4.2 Second Research Question: Across the Four Groups of Participants, What Are the Relationships among Reading Comprehension Test Performance, Character-Recognition Test Performance, and Test-Taker Perception of Strategy-Use While Taking the Reading Comprehension Test?

To answer the second research question, initial confirmatory SEM was conducted without EFA, because the three measurement models were already established in previous studies (for the measurement models of the reading comprehension test and the strategy-use survey: Phakiti, 2007; Purpura, 1997; for the measurement model of the character-recognition test: Li et al., 2002; Tong, 2008). The software used to conduct SEM was IBM SPSS AMOS 22. The subsequent EFA that was implemented to generate suggestions for model re-specification used IBM SPSS Statistics 21.

4.2.1 Initial confirmatory SEM. Prior to conducting the initial SEM, the multivariate normality assumption was checked and met (Mardia's = 1.84), and the number of freely-estimated parameter (10) was smaller than 12(12 +1)/2, ensuring that the degree of freedom of the initial SEM model was larger than zero. All factor loadings of the error and residual terms were constrained at one. Next, the factor loadings of the cloze subtest, the ortho-phonological subtest, the comprehending strategies, and the evaluating strategies were set at one because their reliability coefficients were the highest in their corresponding measurement models.

4.2.1.1 Results of parameter estimation. ML estimation was adopted as recommended by Mueller and Hancock (2010), and the standardized parameters can be seen in Figure 9 and Table 20. All parameters, except for the factor loading from cognitive strategies onto the reading-test scores ($p = 0.08$), were significant at $p < 0.05$.

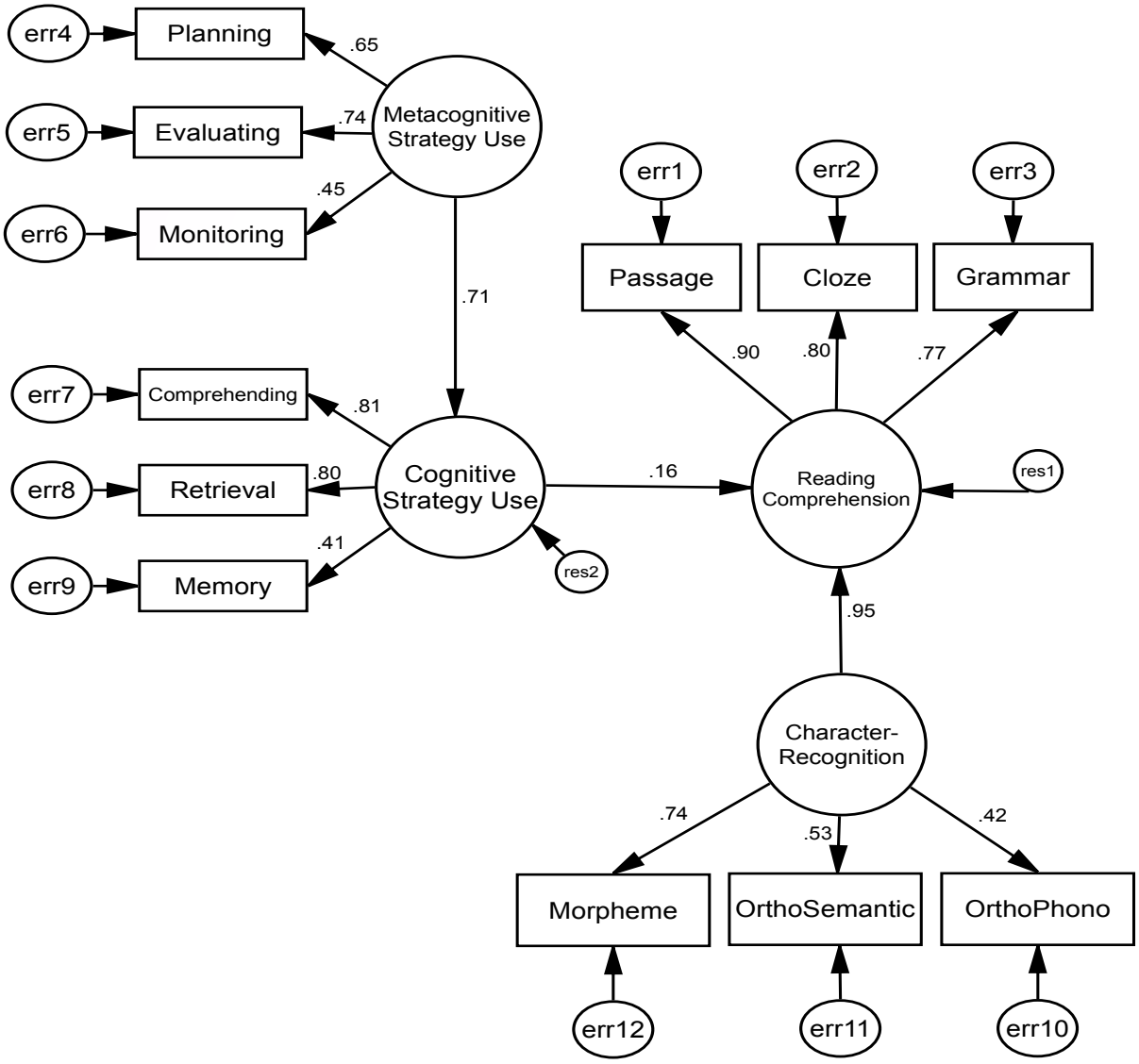


Figure 9. The hypothesized model examining the effects of character-recognition ability and strategy-use on reading comprehension ability. All path parameters, other than the one from cognitive-strategy use to reading comprehension ($p = 0.08$), are significant at $p < 0.05$.

Table 20

Parameter Estimates of the Hypothesized SEM Model

Parameter			Unstandardized		Standardized Estimate	*p*
			Estimate	S.E.		
Cognitive	←	Metacognitive	1.00	NA	0.71	NA
Reading.	←	Cognitive	0.11	0.06	0.16	0.08
Reading.	←	Character-Recognition	2.04	0.60	0.95	< 0.01
Passage	←	Reading.	0.91	0.10	0.90	< 0.01
Cloze	←	Reading.	1.00	NA	0.80	NA
Planning	←	Metacognitive	0.80	0.16	0.65	< 0.01
Memory	←	Cognitive	0.40	0.11	0.41	< 0.01
Retrieval	←	Cognitive	0.69	0.11	0.80	< 0.01
Comp.	←	Cognitive	1.00	NA	0.81	NA
OrthoS	←	Character-Recognition	1.31	0.42	0.53	< 0.01
Grammar	←	Reading	0.72	0.09	0.77	< 0.01
OrthoP	←	Character-Recognition	1.00	NA	0.42	NA
Morpheme	←	Character-Recognition	1.60	0.45	0.74	< 0.01
Eval.	←	Metacognitive	1.00	NA	0.74	NA
Monitor.	←	Metacognitive	0.63	0.18	0.45	< 0.01

Note: Reading. = reading comprehension
Comp. = comprehending
OrthoS = ortho-semantic task
OrthoP = ortho-phonological task
Eval. = evaluating
Monitor. = monitoring

The measurement model of the reading test, as assessed by the grammar, cloze, and passage-comprehension subtests, was designed to tap into higher-level comprehension ability. In other words, performance on the reading test represents a limited set of higher-level skills related to orchestrating decoded information from the syntactic, semantic, and discourse levels. The factor loadings of the measurement model show that reading comprehension ability was adequately assessed by the three subtests: 0.77 (grammar), 0.80 (cloze), and 0.90 (passage comprehension). These parameters indicate the covariance between the reading comprehension ability and the

three subtests, R^2, ranged from 0.59 for the grammar subtest to 0.81 for the passage-comprehension subtest.

The measurement model of the character-recognition test, as assessed by the ortho-phonological subtest, the ortho-semantic subtest, and the morpheme subtest, was designed to tap into character-recognition ability. Test performance on the character-recognition test represents lower-level skills related to decoding phonological and semantic information at the sub-character level, and semantic information at the character level. The factor loadings suggest that the character-recognition test was well measured by the morpheme subtest (0.73), but the factor loadings of the ortho-phonological and the ortho-semantic subtests were moderate (at 0.42 and 0.53, respectively). When the factor loadings were transferred to R^2, the effect sizes of character-recognition skills on the three subtests ranged from 0.18 for the ortho-phonological subtest to 0.53 for the morpheme subtest. In addition, the hypothesized SEM model suggests a high correlation between comprehension ability and character-recognition ability, which in turn could imply that reading ability involves varying degrees of both lower- and higher-level processing skills.

Based on the strategy-use model first proposed by Purpura (1997) and later adjusted by Phakiti (2008), strategy use can be divided into cognitive and metacognitive types. While cognitive strategies directly facilitate reading comprehension, metacognitive ones regulate the use and effectiveness of cognitive strategies. When tested in the hypothesized SEM model, cognitive-strategy use was well explained by the combination of comprehending (0.81), retrieval (0.80), and memory (0.41) strategies, providing R^2 values ranging from 0.17 to 0.66. The model also suggested that metacognitive-strategy use was well explained by the combination of

planning (0.62), evaluating (0.74), and monitoring (0.45) strategies, with R^2 values ranging from 0.21 to 0.55.

In the measurement models of cognitive-strategy use and metacognitive-strategy use, the weakest relationships were approximately 0.40, memory (0.41) in the measurement model of cognitive-strategy use and monitoring (0.45) strategies in the measurement model of metacognitive-strategy use. Although their relationships with the corresponding latent variables were weak, both paths were still statistically significant. Also, the results showed that metacognitive-strategy use had a direct and positive impact on cognitive-strategy use, supporting the findings of Phakiti (2008) and Purpura (1997).

An important goal of this dissertation is to examine the role of strategy use in L2 Chinese reading, but the path cognitive-strategy use → reading comprehension in Figure 9 (r = 0.16) was not statistically significant in the initial SEM model, indicating a need to re-specify the model. This finding is curious, in that it may suggest the participants' cognitive-strategy use only weakly correlated to their reading-comprehension ability, and that such a correlation could be observed only by chance. Other paths between strategy use and L2 Chinese reading skills had to be explored, because without a statistically significant path connecting cognitive-strategy use to reading comprehension, the constructs in Figure 9 would remain divided into two separate parts, i.e., one for strategy use and the other for L2 Chinese reading skills. Additionally, the nonsignificant path may be a function of the present study's small sample size.

4.2.1.2 Results of model fit. Model-fit indices are summarized in Table 21. The χ^2 value suggests that the data is significantly different from the hypothesized model. In addition, the absolute, parsimonious, and incremental fit indices suggest that model fit is less than ideal. First, the higher-than-ideal SRMR estimates suggest a large discrepancy between the implied

covariance in the hypothesized model and the observed covariance in the data. Secondly, after adjusting for model complexity, the discrepancy between hypothesized covariance and observed covariance (as indicated by RMSEA) is still larger than ideal; and the 90% confidence interval of RMSEA is larger than 0.05, showing that the RMSEA estimate is not ideally precise, either. Thirdly, the lower-than-ideal CFI and TLI show that the hypothesized model is not significantly better than a null model in terms of model-data fit. In sum, although the model-fit indices were less than ideal, model fit appeared to be acceptable in light of the small sample size of the study.

Table 21.

Model-Fit Indices of the Hypothesized SEM Model

	χ^2	*df*	SRMR	RMSEA	RMSEA 90% CI		CFI	TLI
					Low	High		
Estimate	85.10*	52	0.12	0.09	0.05	0.12	0.90	0.87

* $p < 0.05$

4.2.1.3 Modification identification. As indicated in Table 22, the recommendations for re-specification of the SEM model included only one parameter to add, representing a covariance between Error 4 and Error 8. According to Byrne's (2010) three recommendations for model re-specification discussed above in section 3.6.4.4, i.e., that any modification be theoretically supported, unlikely to lead to overfitting, and a substantial improvement, the suggested parameter appears to be of little concern, and was ignored in the revised SEM model.

Table 22

Modification Index for the Hypothesized SEM Model

Parameter to Add	Modification Index	Parameter Change
Err4 ←→ Err8	6.32	-0.02

4.2.2 Model re-specification. Although model-fit indices were acceptable in relation to the sample size, the path cognitive-strategy use → reading comprehension was of concern due to its non-significant value and its theoretical importance in the current study. To explore other possible paths between strategy-use and L2 Chinese reading skills, I re-examined the relevant literature discussed in Chapter 2, above. Li (1998) found that CFLLs actively adopted strategies for recognizing unfamiliar characters, or morphemes; Everson and Ke (1997) and Lee-Thompson (2008) found that CFLLs regularly monitored and evaluated their comprehension quality, and used reading strategies to facilitate both lower- and higher-level reading processes, but that most of their strategies focused on deciphering unfamiliar Chinese words and characters; and Lee-Thompson (2008) and Shen (2013) found that CFLLs adopted both metacognitive and cognitive strategies in their character-recognition processes. Taken together, these studies' results suggested that my model re-specification should examine the relationship between strategy use and character recognition, including both metacognitive and cognitive strategies for character recognition. However, the abovementioned studies were mainly based on frequency data collected through verbal reports, and their strategy categories were mostly a combination of all six of the strategy types discussed in this dissertation, leading to challenges in identifying the main path(s) connecting strategy use and reading processes that would be suitable to the SEM model adopted here. To narrow down which strategy type(s) would best relate to character recognition, I first examined the complete correlation matrix (Table 23) and then conducted a EFA.

Table 23

Correlations between Variables

	1	2	3	4	5	6	7	8	9	10	11	12
1. Grammar		0.623*	0.678*	0.046	-0.111	0.147	0.026	-0.121	-0.016	0.414*	0.330*	0.529*
2. Cloze			0.719*	-0.054	-0.140	0.058	-0.056	-0.289*	-0.034	0.333*	0.353*	0.543*
3. Passage				0.009	-0.144	0.123	-0.243	-0.278*	-0.069	0.360*	0.482*	0.627*
4. Comp.					0.244*	0.584*	0.192	0.141	0.306*	0.040	0.001	-0.186
5. Memory						0.269*	0.197	0.233*	0.363*	0.025	-0.149	-0.182
6. Retrieval							0.133	0.248*	0.411*	0.049	0.027	-0.102
7. Plan.								0.276*	0.598*	-0.051	-0.160	-0.322*
8. Monitor.									0.380*	-0.034	-0.082	-0.253*
9. Eval.										-0.060	-0.011	-0.090
10. OrthoP											0.236*	0.243*
11. OrthoS												0.438*
12. Morpheme												

Note: Comp. = comprehending
Plan. = planning
Monitor. = monitoring
Eval. = evaluating
OrthoS = ortho-semantic task
OrthoP = ortho-phonological task
*$p < 0.05$

As Table 23 indicates, all the correlations between the reading and character-recognition subtests were statistically significant, except those related to strategy-use variables. Taken together with the high factor loading of the character-recognition → reading path in the initial SEM model (r = 0.95 in Figure 9), this suggests that the reading processes of reading comprehension and character recognition are significantly correlated to each other, although the magnitudes of such correlations varied considerably. These statistically significant correlations could indicate certain trends. First, the higher-level skills assessed by the three reading subtests may belong to the same latent variable: reading-comprehension ability. Second, though significant correlations were found between all the lower-level skills assessed by the three character-recognition subtests, the highest correlation was between ortho-semantics and morpheme skills (r = 0.44). The medium size of this correlation may suggest that, for participants in this study, the two lower-level skills were more closely correlated, perhaps because both process semantic information; whereas ortho-phonology skills' relationships with the two lower-level semantics skills were significant but smaller (r = 0.24 with ortho-semantics and r = 0.24 with morpheme skills). Third, all the correlations between reading-comprehension and character-recognition skills were statistically significant, and the correlations between the three reading subtests and morpheme skill were higher than those between the same subtests and the other two lower-level skills. The stronger correlations between morpheme skill and higher-level skills may suggest that morpheme skill played a key role in the participants' connection of higher- to lower-level skills.

With regard to the relationships between strategy use and other reading skills, the trends in the correlations were less clear. First, among the correlations between strategy types and other reading skills, only four correlations were statistically significant. These were monitoring

strategies with two higher-level skills, and planning and monitoring strategies with morpheme skills. The absolute value of the correlations between strategy use and higher-level reading skills was approximately 0.28, while the absolute value of the correlations between strategy use and lower-level reading skills ranged from 0.025 to 0.32. As such, it appeared that strategy use may be more strongly correlated with morpheme skill than with reading-comprehension ability.

Second, among the six strategy types, planning strategies correlated more closely with the other two metacognitive strategies (r = 0.28 with monitoring, and r = 0.60 with evaluating) than with any of the three cognitive strategies. Planning strategies are used to map out the necessary steps for completing current reading tasks, to clarify reading goals, and to scan through reading tests before starting them. In other words, planning strategies are activated for the purpose of coordinating other reading strategies as well as cognitive resources. However, judging from planning strategies' lower correlations with cognitive strategies and higher correlations with the other metacognitive strategies (see Table 23), it is possible that the participants in the current study used planning strategies for coordinating metacognitive strategies more than for coordinating cognitive ones.

Third, the correlations between cognitive strategies and evaluating strategies (ranging from 0.31 to 0.41) were stronger than those between cognitive strategies and the other metacognitive strategies (ranging from 0.14 to 0.25). Additionally, the correlations of evaluating strategies with metacognitive strategies (r = 0.60 with planning, and r = 0.38 with monitoring) were stronger than the correlation between the other two metacognitive strategies (r = 0.28). Evaluating strategies' stronger correlations may suggest that they functioned to connect metacognitive and cognitive strategies among this study's participants.

Fourth, strategy use tended to correlate better with passage-comprehension and morpheme subtests than with the other four subtests. Moreover, among the significant correlations between strategy use and subtests other than passage-comprehension and morpheme, some of the significant correlations were negative, and these correlations were related to planning and monitoring strategies. This might indicate that planning and monitoring strategies were activated to compensate for interrupted reading processing.

The trends in Table 23 can be summarized as follows: 1) morpheme skills may be crucial for connecting the lower- and higher-level processes; 2) the correlations between morpheme skill and strategy use, and monitoring and evaluating strategies in particular, are stronger than those between higher-level processes and strategy use, especially monitoring; 3) planning may coordinate more metacognitive strategies than cognitive strategies; and 4) evaluating strategies may be responsible for connecting metacognitive and cognitive strategies.

Next, a EFA was conducted to examine the correlational trends in Table 23 and to explore potential paths for a modified SEM model. The number of extracted factors was decided based on their eigenvalues and the scree plot. As can be seen in the scree plot in Figure 10, there were three factors with eigenvalues larger than one. However, the scree plot indicates that from Factor 3 onward, no substantial variance was added. In other words, the two criteria suggest that the number of extracted factors should be two or three.

In the initial SEM model, correlations among latent variables were assumed; thus, direct oblimin rotation, using the default delta value of zero, was conducted. When examining the pattern matrix, the factor loadings (shown in Table 24) indicate that a three-factor solution would be a better fit than a two-factor solution in light of the theoretical considerations of the study. Therefore, the 12 variables clustered into three factors, representing reading ability, cognitive-

strategy use, and metacognitive-strategy use. Much as in the results of the initial SEM, the reading-test and character-recognition test variables loaded highly on Factor 1; cognitive strategies loaded highly on Factor 2; and metacognitive strategies loaded highly on Factor 3. However, complicated factor loadings in Factors 2 and 3 were also found. Namely, the factor loadings of memory strategies, as a variable loading on Factor 2, on Factors 2 and 3 were 0.28 and 0.20, respectively; and the factor loadings of monitoring strategies, as a variable loading on Factor 3, on Factors 2 and 3 were 0.19 and 0.29, respectively. The factor-correlation matrix in Table 25 indicates that Factor 2 and 3 have a correlation of 0.43, and this may explain the complicated factor loadings found for memory and monitoring strategies.

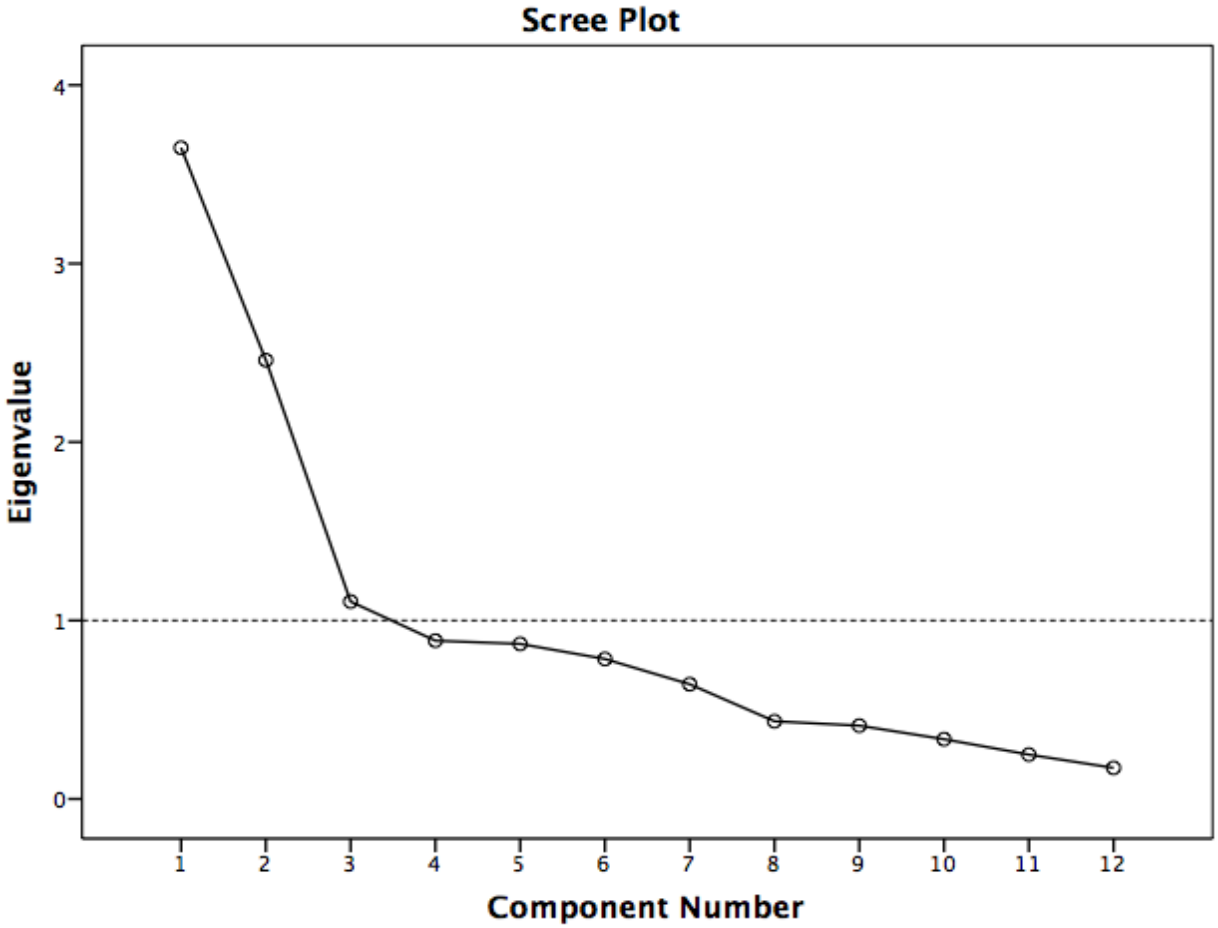


Figure 10. Scree plot of the EFA

Table 24

Rotated Factor Loadings

	Factor			Communalities
	1	2	3	
Grammar	**0.81**	0.03	0.15	0.62
Cloze	**0.81**	-0.07	0.10	0.64
Passage	**0.88**	0.11	-0.11	0.83
Comp.	0.00	**0.64**	0.01	0.42
Memory	-0.12	**0.28**	0.20	0.20
Retrieval	0.11	**0.92**	-0.07	0.80
Planning	0.01	-0.14	**0.89**	0.70
Monitoring	-0.19	0.19	**0.29**	0.23
Evaluating	0.10	0.22	**0.68**	0.62
OrthoP.	**0.43**	0.05	0.02	0.18
OrthoS.	**0.49**	0.03	-0.06	0.26
Morpheme	**0.68**	-0.13	-0.10	0.54
% Variance Explained	0.27	0.17	0.06	**0.50**

Comp. = comprehending strategies
OrthoP = ortho-phonological subtest
OrthoS = ortho-semantic subtest

Table 25

Correlation between Factors

Factor	1	2	3
1	1	-0.05	-0.23
2		1	0.43
3			1

It also emerged that the correlation between Factors 1 and 3 was higher than that between Factors 1 and 2, even though metacognitive strategies indirectly regulate reading comprehension through the direct facilitation of cognitive-strategy use on reading comprehension. In parallel to the correlational trends observed in Table 23, the correlation between Factors 1 and 3 was higher

than the correlation between Factors 1 and 2, which may suggest that metacognitive strategies were activated by interrupted reading processes. Moreover, when examining the unique variances between Factor 1 and the three metacognitive-strategy types, the factor loading of monitoring strategies was the highest among the three. On the other side of the coin, based on the unique variances between Factor 3 and the six reading skills, its correlations with character recognition and with reading comprehension were similar, although somewhat higher with the later.

Therefore, the paths monitoring → reading comprehension and monitoring → character recognition were separately added to the initial SEM model, and the two revised SEM models were assessed simultaneously. Then, another revised SEM model with both these additional paths was assessed. The evaluation of the three resulting revised SEM models followed the steps listed in Table 26. The models with one added path apiece were analyzed and compared against each other; and the better-fitting of the two (i.e., Model 1b) was further compared against the revised SEM model that featured both of these additional paths.

Table 26

Model-Fit Indices of the Hypothesized Model and Revised Models

Step	Parameter to Add	χ^2	*df*	$\Delta\chi^2$	Δdf	SRMR	RMSEA	CFI	TLI
0	Initial Model	85.10*	52	NA	NA	0.11	0.12	0.90	0.87
1a	Monitor. → Reading.	79.84*	51	5.26*	1	0.11	0.08	0.91	0.88
1b	Monitor. → C.R.	76.69*	51	8.41*	1	0.09	0.08	0.92	0.90
2	Monitor. → Reading. Monitor. → C.R.	75.79*	50	4.05* 0.90	1	0.09	0.08	0.92	0.89

* $p < 0.05$

Note: Monitor. = Monitoring
Reading = Reading Comprehension
C.R. = Character Recognition

After adding either of the two paths, χ^2 values changed significantly, and other model-fit indices slightly improved as well. Such indices further suggested that Model 1b was better than Model 1a, and thus Model 1b was chosen to compare against Model 2. However, in the latter comparison, the difference in χ^2 between Model 1b and Model 2 was not significant and other model-fit indices did not improve, either. Therefore, Model 1b appeared to be the best-fitting revised model among the three models due to the considerations of model parsimony and over-specification.[8]

[8] In Model 2, the path monitoring → character recognition (r = -0.27) and the path monitoring → reading comprehension (r = -0.27) were not statistically significant, although the path cognitive-strategy use → reading comprehension became statistically significant (r = 0.24). In other words, the parameter estimates in Model 1b were also better than those in Model 2, although the model-fit indices between the two models were not significantly different.

4.2.3 Revised SEM.

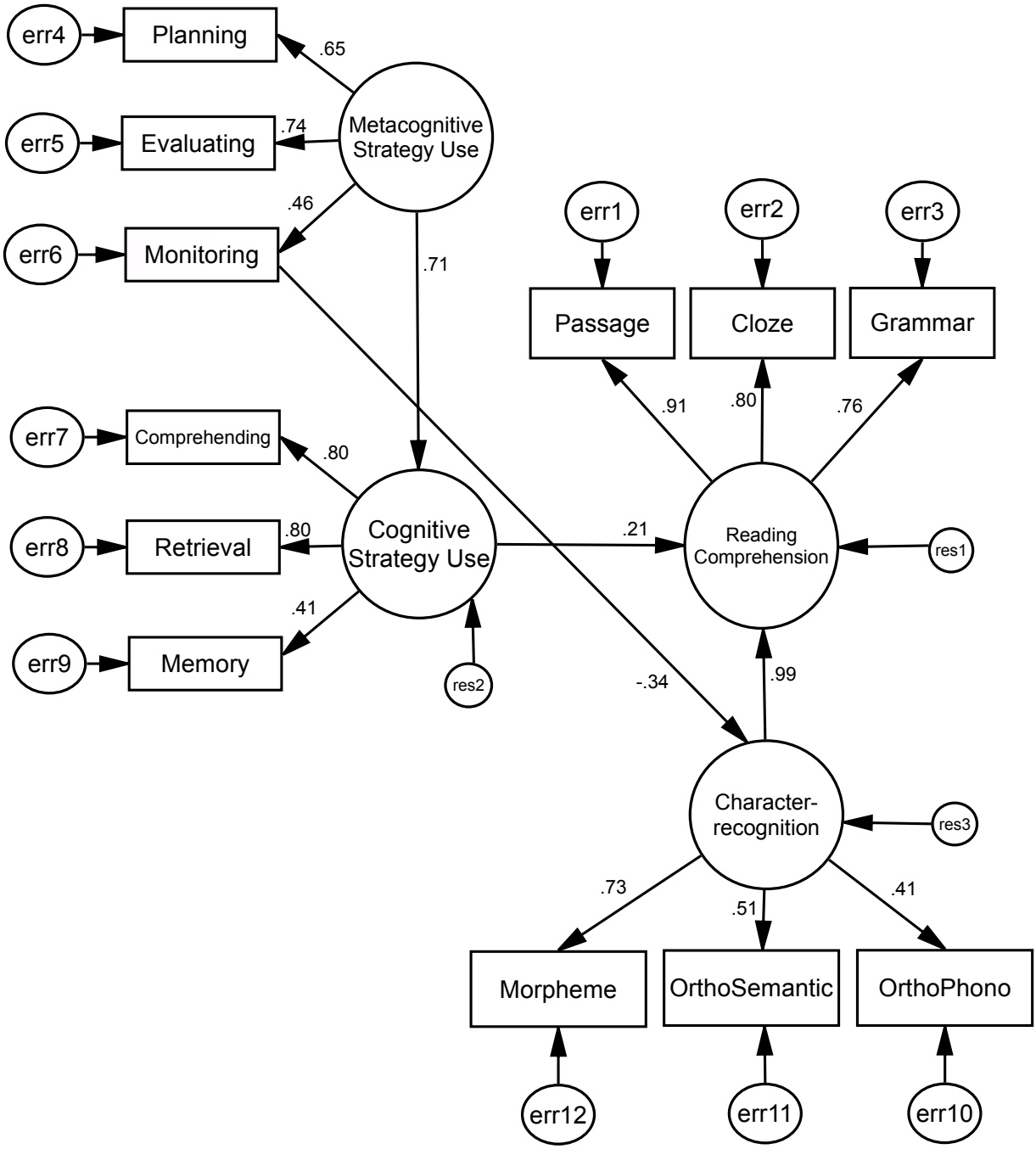


Figure 11. The revised model examining the effects of character-recognition skills and strategy use on reading comprehension ability. All path parameters are significant at $p < 0.05$.

4.2.3.1 Results of parameter estimation. Table 27 contains a summary of the parameter estimates for the revised model, Model 1b.

Table 27

Parameter Estimates of the Revised SEM Model

Parameter			Unstandardized		Standardized	*p*
			Estimate	S.E.	Estimate	
Monitoring	←	Metacognitive	0.64	0.18	0.46	< 0.01
Cognitive	←	Metacognitive	1.00	NA	0.71	NA
Character-Recognition	←	Monitoring	-0.10	0.04	-0.34	0.02
Reading.	←	Cognitive	0.14	0.06	0.21	0.02
Reading.	←	Character-Recognition	2.17	0.64	0.99	< 0.01
Passage	←	Reading.	0.91	0.10	0.91	< 0.01
Cloze	←	Reading.	1.00	N/A	0.80	NA
Planning	←	Metacognitive	0.80	0.16	0.65	< 0.01
Retrieval	←	Cognitive	0.70	0.11	0.80	< 0.01
Comp.	←	Cognitive	1.00	N/A	0.80	NA
OrthoS.	←	Character-Recognition	1.29	0.42	0.51	< 0.01
Grammar	←	Reading	0.71	0.09	0.76	< 0.01
OrthoP.	←	Character-Recognition	1.00	N/A	0.41	NA
Morpheme	←	Character-Recognition	1.61	0.45	0.73	< 0.01
Evaluating	←	Metacognitive	1.00	N/A	0.74	NA
Memory	←	Cognitive	0.40	0.12	0.41	< 0.01

Note: Reading. = Reading Comprehension
Comp. = comprehending strategies
OrthoP = ortho-phonological subtest
OrthoS = ortho-semantic subtest

The magnitudes of these estimates were similar to those in the model shown in Figure 11; but a key difference between the two models is that all the parameters in the revised model were statistically significant, $p < 0.05$, after the inclusion of the monitoring → character-recognition path. The factor loading the cognitive-strategy use → reading comprehension path became statistically significant and slightly higher (0.21 versus 0.16). Moreover, the factor loading of the newly added path was -0.34, indicating that there was a negative correlation between monitoring strategies and fluent character recognition. In other words, it is likely that readers activated other

metacognitive strategies when monitoring strategies detected interrupted character recognition. Furthermore, based on the squared multiple correlations for the latent variable of reading comprehension, 99% of the variance in comprehension ability was directly and indirectly accounted for by character-recognition ability, cognitive-strategy use, metacognitive-strategy use, and monitoring-strategy use.

4.2.3.2 Results of model fit. The model-fit indices in Table 28 suggest that the fit of the revised model is reasonable, considering the relatively small sample size. The 90% confidence interval indicates that the confidence interval of RMSEA estimate is larger-than-ideal and the RMSEA estimate can be as high as 0.11, which might also be due to the limitations imposed by the smallness of the sample.

Table 28

Model-Fit Indices of the Revised SEM Model 1b

χ^2	*df*	SRMR	RMSEA	RMSEA 90%CI		CFI	TLI
				Low	High		
76.69*	51	0.09	0.08	0.04	0.11	0.92	0.90

* $p<0.05$

4.2.3.3 Modification identification. Table 29 shows a proposed modification provided by the SEM analysis, representing error covariance between Error 4 and Error 8. However, this was ignored due to its low importance to the second research question.

Table 29

Modification Index for the Revised Model

Parameter to Add	Modification Index	Parameter Change
Err4 <-> Err8	6.65	-0.02

In answer to the second research question, the revised SEM model presented in Figure 11 provided strong evidence in favor of the interactive model of L2 Chinese reading. First, the reading comprehension ability measured by the reading test was highly correlated with the character-recognition ability measured by the character-recognition test (0.99), suggesting that the two latent variables were closely related and might belong to a higher-order construct.

Metacognitive-strategy use and cognitive-strategy use were also highly correlated (0.71), and yet still clustered into two different components. The divergence but relatedness of these two latent variables appears to provide empirical evidence that metacognitive strategies regulate the use of cognitive ones. In other words, readers consciously monitor their affective status and comprehension quality, and select the cognitive strategies most suitable to facilitating their higher- and lower-level reading processes.

The relationships between strategy use and reading processing were more complicated. The revised SEM model showed that cognitive-strategy use had a direct, positive impact (0.21) on the participants' reading-test performance; however, contrary to my hypothesis, monitoring strategy also exerted a direct, negative impact (-0.34) on character-recognition ability. Moreover, the two paths connecting strategy use to other reading processes indicated that such relationships were more complicated than the hypothesized SEM model proposed in Figure 1, above.

4.3 Third research Question: Across The Four Groups of Participants, To What Extent Does Character-Recognition Versus Strategy-Use Factors Affect Reading-Test Performance?

Examining the structural model in the revised SEM model, both character-recognition ability and strategy use significantly affected reading-test performance. The parameters showed that correlation between character-recognition score and reading-test score was 0.99, which

contrasted sharply with the correlation between cognitive-strategy use and reading-test score (0.21). However, perceived monitoring strategy use, as a metacognitive strategy, had an additional direct and negative impact (-0.34) on reading-test performance. The direct effect of cognitive-strategy use on comprehension processing, although small, significantly influenced reading-test performance. Also, metacognitive-strategy use was highly correlated with cognitive-strategy use (0.71); and the use of a monitoring strategy, through its direct effect on the character-recognition test, indirectly influenced performance on the reading test.

Compared with the larger effect of character-recognition skills, the effect of strategy use on reading-test scores was relatively small; nevertheless, it was statistically significant. Strategy-use also exhibited a more complicated effect on the reading test. First, in terms of direct effects, cognitive-strategy use had a positive effect, and it appears that the more readers perceived themselves to be using cognitive strategies, the better their performance on the reading test would be. Second, regarding an indirect effect, metacognitive-strategy use had a positive effect on reading-test scores through its regulation of cognitive-strategy use. In other words, the higher one's perceived use of metacognitive strategies was, the better one performed on the reading test. Third, as another indirect effect on the reading test performance, the perceived adoption of a monitoring strategy had a negative effect on character-recognition test outcomes, meaning that the higher one's perceived use of monitoring strategies was, the worse one would perform on the character-recognition test.

These direct and indirect effects of strategy use may be explained, first through the negative effect of monitoring strategies. Such strategies include pacing oneself while taking the test, and maintaining awareness of one's concentration status and quality of comprehension. Monitoring strategies consistently examining affective and comprehension status, and when

confronting unfamiliar Chinese characters, and being distracted and stressful, monitoring strategies may activate other metacognitive strategies and choose the most appropriate cognitive strategies to facilitate comprehension ability for compensating the interrupted character-recognition processes.

4.4 Summary of Results for Research Questions

The aim of this study was to investigate an interactive model of L2 Chinese reading while taking account of language background as a confounding variable. In this chapter, the composite scores for the reading test, the character-recognition test, and the strategy-use survey were used to examine an interactive model of L2 Chinese reading. Several variables from lower-level and higher-level processing skills were chosen. A group of lower-level skill variables was included to represent character-recognition skills, and three groups of higher-level processing variables were chosen to represent background knowledge, reading comprehension skills, and strategy use. Background knowledge was examined through participants' language background in terms of heritage or foreign-language learning experience. As a confounding and nesting variable of the interactive model, the influences of language background on other variables were investigated first, through Research Question 1. Later, all the participants were clustered into a single group, to allow for investigation of common constructs of L2 Chinese reading that may be shared by all language learners.

4.4.1 Research Question 1: To what extent can language background explain the variance in performance on the reading test, the character-recognition test, and the strategy-use survey?

Originally there were three language backgrounds included in the study: CFLLs, Man-HLLs, and Can-HLLs. However, due to the limitations of convenience sampling, a subgroup of

CFLLs was singled out after examination of descriptive statistics, leading to CFLLs being divided into Singaporean CMTLLs and CFLLs. As a result, four language backgrounds were included in the efforts to answer the first research question. The results of three profile analyses showed that language background only explained 18% of reading-test variance, 9% of character-recognition test variance, and 5% of strategy-use survey variance, with only the first one result being statistically significant. The interaction between language background and reading skills, meanwhile, explained 8% of reading-test variance, 7% of character-recognition test variance, and 6% of strategy-use survey variance. In sum, the effect size of language background was less than 20% of the reading variance.

4.4.2 Research Question 2: Across the four groups of participants, what are the relationships among 1) reading-test performance, 2) character-recognition test performance, and 3) test-taker perception of strategy-use while taking the reading test?

The primary constructs included in the hypothesized SEM model were revised based on relevant research findings and the results of intercorrelations between variables and EFA. In the revised SEM model, all the specified paths were statistically significant. In addition, performance on the reading test was largely predicted by performance on the character-recognition test, and moderately predicted by cognitive-strategy use. Moreover, metacognitive and monitoring-strategy use indirectly, but significantly, predicted performance on the reading test as well. The relationship between strategy use and other reading processes is more complicated, in terms of the signs of the two factor loadings: a positive and direct effect of cognitive-strategy use on reading-test performance, and a negative and indirect effect of monitoring strategies on such performance (through the character-recognition test).

4.4.3 Research Question 3: Across the four groups of participants, to what extent did character-recognition versus strategy-use factors affect reading-test performance?

In the revised SEM model, the significant predictors of reading-test performance were character-recognition skills (r = 0.99, direct effect), cognitive-strategy use (r = 0.21, direct effect), metacognitive-strategy use (r = 0.15, indirect effect), and monitoring-strategy use (r = -0.33, indirect effect), together accounting for 99% of the reading-test score variance. Thus, it can be said that character-recognition ability affects reading-test performance to a much larger extent than strategy use does. Additionally, in comparison to the strong and direct effect of character recognition on reading-test performance, the effect of strategy use was subtler and complicated, as indicated by cognitive-strategy use having a direct effect, but metacognitive and monitoring strategies having indirect ones.

CHAPTER 5

DISCUSSION AND CONCLUSIONS

The purposes of this study were to investigate the constructs of L2 Chinese reading, and to give due consideration to Chinese-language learners as active agents rather than passive processors. In order to achieve these two aims, an interactive model of L2 Chinese reading was developed, in which lower-level character-recognition, higher-level comprehension, and higher-level strategy-use skills were measured through a strategy-use survey and the strategy-use survey covered six strategy types. Crucially, the participants' language backgrounds were also included as a background-knowledge variable, which had the potential to confound all the other processing skills.

The study was based on the hypothesis that a portion of the variability in the participants' performance on a test of Chinese reading ability could be explained by character-recognition ability, strategy use, and language background. This chapter first summarizes the key findings and compares them against those of previous studies, and then goes on to discuss their theoretical and pedagogical implications. The discussions of key findings begin with the effect of language background, followed by character-recognition skills, comprehension skills, the relationship between the two aspects of reading skills, and strategy-use, and ends with the full SEM model of the studied processing skills.

Concerning the low reliability coefficients of some subtests, Rasch analyses were conducted in hope to improve the internal consistency by transferring raw scores into true interval variables (Bond & Fox, 2015). After obtaining ability estimates of the 12 variables using the software Facets 3.71.0 (Linacre, 2013), reliability of the 12 subtests increased and the reliability coefficients were around 0.75 (see Appendix C). However, because each subtest had

only a few items (11 or fewer in each case), 135 misfitting ability estimates were generated across the 12 variables, representing 13% of the data. Due to substantial amount of misfitting estimates over small number of items, I chose to use raw data to conduct further analyses, bearing in mind the low reliability of some subtests that this would entail. The descriptive statistics of each group based on Rasch ability estimates are presented in Appendix D.

5.1 The Effect of Language Background on L2 Chinese Reading Skills

It should be noted that due to the small sample size and multiple comparisons employed in this study, the observed power of many of the analyses is low, and may have led to some non-significant results. Notwithstanding such limitations, the results do suggest that language background had a medium effect on reading-test performance (approximately 18%), while its effects on character-recognition test performance and strategy-use survey responses were both small (less than 10%). All three of the abovementioned results echoed the non-significant group differences found in previous studies (Ke, 1998; Koda, Zhang, & Yang, 2008; Lü & Koda, 2011; Xiao, 2008). However, though the effect of language background on lower-level processing was not statistically significant, its effect on higher-level processing was larger and significant, and this has several implications regarding the effect of language background on L2 Chinese reading processing. The effect of language background is discussed according to the following three aspects: 1) its stronger effect on comprehension processing than on character-recognition processes, 2) its weak effect on strategy use, and 3) the social side of the language-background effect observed in Singaporean CMTLLs.

First, comprehension processing may be more sensitive to language background than character-recognition processing is. However, upon closer examination of the reading and character-recognition subtests in Figures 6 and 7, group differences appear to be larger in the

grammar, cloze, and morpheme subtests than in the passage-comprehension, ortho-semantic, and ortho-phonological subtests, suggesting the existence of another interaction effect between language background and reading-processing levels. In other words, language background may have a stronger effect on morpheme, vocabulary and grammar knowledge, than on reading processes at the sub-character and passage levels. which is similar to the HLLs' morphosyntactic advantage summarized by Montrul (2012).

In terms of this seeming uniformity of sub-character processing, the findings of Lü and Koda (2011) may offer a possible explanation. They divided their participants into two groups based on their home language use and home literacy support to examine the effects of these factors on three skills related to character learning: real character naming, real pinyin naming, and pseudo-pinyin naming. The only significant effect these authors found was the positive impact of home language use on real character naming,[9] implying that the benefit of language background is mainly at the character-level, while the more refined subcharacter-level skills may require much longer exposure and/or formal language instruction.

On the other hand, the relatively small group difference I observed in the passage-comprehension subtest may be illuminated by the research of Koda, Zhang and Yang (2008) and Schoonen, Hulstijn and Bossers (1998). Both these studies' findings suggest that L2 passage comprehension involves metacognitive variables that are transferable across languages; and that such metacognitive variables include viewing oneself as a reader, knowledge of text characteristics, knowledge of reading strategies, and knowledge of reading goals and comprehension criteria (Schoonen, Hulstijn, & Bossers, 1998). Among these four metacognitive variables, knowledge of text characteristics is developed last, and is the most important in both

[9] Although the effect of home language use on real character naming was the only statistically significant comparison reported by Lü & Koda (2011), its statistical significance is in question due to the inflated Type-I error rate caused by the 18 t-tests conducted in the study.

L1 and L2 passage comprehension. In the current study, all the participants were all aged 16 or above, and many of them had received some higher education, indicating that they might also have been able to transfer their knowledge of text characteristics from English to Chinese, leading language background to have a smaller effect.

Second, although the present study identified a positive relationship between strategy use and comprehension, the effect of language background on strategy use was not statistically significant: a finding that differs from that of Sung (2011). This discrepancy may be related to the present work limiting strategy use to a particular context, i.e. a test preparation context, whereas Sung applied it to the learning of Chinese in general. As such, the narrower range of strategies employed by participants in the present study may have been less useful for tapping into group differences. Even so, the effect size of language background in the present study was about 6% of the total variance in cognitive- and metacognitive-strategy use, which is roughly comparable to the highest effect of home language on the total variance in metacognitive strategy-use in Sung's (2011) study, i.e., 11%. However, in one of the aforementioned studies (Schoonen, Hulstijn, & Bossers, 1998), knowledge of reading strategies and knowledge of reading goals and comprehension criteria – both of which are categorized as strategy use in the present dissertation – exhibited stable effects on L1 and L2 reading comprehension. This further suggests that strategy use may be transferable from English to Chinese, which in turn could explain the nonsignificance of the effect of language background among my participants.

Third, following the screening of descriptive statistics, Singaporean CMTLLs emerged as a distinctive group. Based on their responses to the language-background survey, most of them identified their parents as native speakers of English, likely implying that their families had been resident in Singapore for more than three generations. In contrast to the findings by Kondo-

Brown (2005), that participants with Japanese-speaking grandparents did not perform significantly differently from their FLL counterparts in a Japanese placement test of an American university, the Singaporean CMTLLs in the current study showed the highest mean scores among the four groups in all the subtests. In addition, these Singaporean CMTLLs reported higher use of the target language with their parents than did the participants with Japanese-speaking grandparents in Kondo-Brown's (2005) study.

The marked differences between the Japanese HLLs studied by Kondo-Brown (2005) and the Singaporean CMTLLs in the current study may be partly explicable by a linguistic-vitality framework (Liu, 2013). When investigating the linguistic vitality of Chinese in the United States, Liu focused on three aspects: capacity, opportunity, and desire. With regard to the first of these, ethnic Chinese in Singapore are required to learn Mandarin at school (Ministry of Education, Singapore, 2016b), and as a result, Singaporean CMTLLs are likely to be more capable of developing their language proficiency through formal and informal education. With regard to the second factor, opportunity, ethnic Chinese represent more than 70% of the population in Singapore (Department of Statistics, Singapore, 2015), and thus Singaporean CMTLLs may have more informal opportunities to develop the target language. And thirdly, their relatively high capacity and greater opportunities may well lead to Singaporean CMTLLs having a stronger desire to develop their Mandarin abilities. In short, the higher linguistic vitality of Mandarin in Singapore – as compared to Mandarin or Japanese in the U.S. – may serve as a protecting factor for bi/multi-lingual competence; and in the absence of formal protection for the social factors encompassed by the linguistic vitality framework (e.g., the Mother Tongue Language Policy in Singapore), it is very likely that HLLs will lose their home languages.

5.2 Character-Recognition Skills in L2 Chinese Reading

In the present study, the participants' character-recognition skills – considered as an aspect of lower-level processing – were measured by the ortho-phonological, ortho-semantic, and morpheme subtests. In the SEM measurement model of character recognition, ortho-phonology, ortho-semantics, and morpheme processing were significantly correlated with character-recognition processing, and this echoes the findings of L1 Chinese reading studies (e.g., Cheung et al., 2007; Li et al., 2002; Shu et al., 2003; Tong, 2008). The results also showed that character-recognition skills were most highly correlated with morphology processing (0.73), followed by ortho-semantics processing (0.54) and ortho-phonology processing (0.42). This high correlation of morphology processing with character-recognition was also similar to the findings of L1 Chinese reading studies (Hu, 2013; Ku, 2000, Li et al., 2002; Tong, 2008), and suggests that morphology processing may be a more sensitive indicator of character-recognition ability than sub-character processing skills are – especially in light of Tong's (2008) finding that the importance of morphology processing increases as L1 reading ability develops. Moreover, my participants' morpheme skill was significantly correlated with their other character-recognition and comprehension skills (see Table 23), and its high factor loading in the EFA results (Table 23) indicates its crucial role in L2 Chinese reading. In light of the beneficial effect of home language use on real character naming identified by Lü and Koda (2011), it is reasonable to argue that morpheme skill plays a more crucial role than ortho-phonology skill in transforming knowledge of oral Chinese into reading ability.

Moreover, Tong (2008) found that sub-character processing emerges as an individual latent variable for beginning L1 readers, and then clusters with morpheme processing for advanced L1 readers. Tong explained that the emergence of sub-character-processing may be

due to an instruction effect. Specifically, when being instructed by compound characters comprising of ortho-phonological and ortho-semantic components with reliable phonological and semantic cues, beginning L1 readers learn to actively process sub-character information. The convergence with morpheme processing may be related to later instruction covering compound characters with less reliable ortho-phonological and ortho-semantic components, which could lead a student to depend less on sub-character processing. The lower factor loadings of the two types of sub-character processing identified in this study may also reflect an instruction effect, insofar as most Chinese characters in textbooks do not have reliable sub-character cues, and instructors may focus on whole characters more than on sub-character components.

When comparing the factor loadings of ortho-phonology and ortho-semantics, the latter's higher correlation with the latent variable of character-recognition ability was also reported in studies of L1 Chinese reading (Cheung et al., 2007; Tong, 2008). Cheung and her colleagues found that ortho-phonology and ortho-semantics skills shared 10% and 12%, respectively, of the total variance in reading comprehension among L1 readers. In the current study, the shared variances between ortho-phonology processing and passage comprehension were 13%, and between ortho-semantics processing and passage-comprehension, 23%. This indicates that processing the sub-character level of semantic information may share more cognitive mechanisms with general-comprehension processing than processing the sub-character level of phonemic information.

In Tong's (2008) study, the factor loading of ortho-phonology skills starts from 0.51 for L1 pre-readers and drops to 0.35 for advanced L1 readers, whereas the factor loading of ortho-semantics skills remains stable at around 0.50 across the three examined age groups. Similarly, in the present study, ortho-phonology and ortho-semantics skills shared 13% and 23%,

respectively, of the total variance with the passage-comprehension subtest; and the factor loadings of ortho-phonology and ortho-semantics skills were 0.41 and 0.51, respectively. The similarities between L1 readers and the participants in the current study could suggest that, like L1 Chinese children, Chinese-language learners are able to utilize ortho-phonological and ortho-semantic information to learn Chinese characters.

Scholars investigating reading development in the context of alphabetic orthographies have identified the decoding of ortho-phonological information as crucial (Morais, 1999; Stanovich, 2000). However, studies of bilingual Chinese children (Cheung et al., 2007; Chow, 2014; Hu 2013) suggest that ortho-phonology processing is less important in Chinese than in English, and that when reading in Chinese, it is also less important than ortho-semantics processing. In addition, the CFLLs studied by Shen (2010) had more difficulty connecting graphemes with phonemes than connecting graphemes with meanings. Along the same line, English-dominant participants in the same study exhibited lower loadings of ortho-phonology processing than of ortho-semantics processing: indicating that English-dominant readers are able to adjust their lower-level processing skills to an orthography very different from that of their dominant language. This processing adjustment would tend to support the concept of multi-competence proposed by Cook (2002): i.e., that bi/multi-linguals develop different skills from their monolingual peers, and that their adaptability to various languages suggests a powerful cognitive ability to choose the most efficient way to process languages, rather than passively relying on the cognitive ability developed through L1 acquisition (see also Cook, 2015).

5.3 Comprehension Skills in L2 Chinese Reading

Comprehension skills, as an aspect of higher-level processing, were measured by the grammar, cloze, and passage-comprehension subtests, targeting the lexical, syntactical, and

discourse levels of comprehension. In the measurement model of reading, the passage-comprehension subtest had the highest factor loading, although the other two subtests also had relatively high correlations with the latent variable of reading comprehension ability. These high factor loadings suggest that the comprehension skills elicited by the three subtests were highly correlated with reading-comprehension ability as a construct: a similar result to Purpura's (1999) regarding the construct of reading ability.

5.4 The Relationship between Character-Recognition and Comprehension Skills

The extremely high factor loadings between character-recognition skills and reading-comprehension skills ($r = 0.99$) in the revised SEM model shown in Figure 11 suggests that these two aspects of L2 Chinese-reading processing are likely to belong to the same higher-order construct. It should be noted here that the unidirectional arrow from character recognition to reading comprehension was established for a statistical reason, so reading comprehension could serve as the dependent variable in the SEM model. In practice, the relationship between the two constructs is bi-directional, meaning that decoded information by the comprehension and character-recognition processes feed into each other. Accordingly, in the structural model, the correlation between comprehension and character-recognition abilities being the highest one suggests that most reading acts would be related to interactions between these two constructs. For example, adequate reading comprehension can be built by piecing together information decoded via character-recognition. At the same time, if a person has sufficient comprehension at the discourse level, this person can guess the meaning of unfamiliar words through accessing information at the morpheme level.

When investigating the relationship between L2 English reading-test performance and strategy use, Purpura (1999) divided reading into reading ability and grammar ability, with the

former measured by passage-comprehension and cloze tests, and the latter by grammar, vocabulary, word-formation, and sentence-formation tests. Purpura found a significant correlation, as high as 0.99, between these two sets of abilities. In the current study, grammar was measured as an aspect of higher-level processing, and no specific subtest was designed to measure word-level processing, although the cloze subtest arguably included aspects of it. On the other hand, due to the importance of morphemes in Chinese reading, a construct of character-recognition was included to tap into this lower-level processing realm. The high correlation between lower-level and higher-level reading processing found in Purpura (1999) and the current study may yet be observed in other L2 reading contexts.

Often, when scholars describe reading in Chinese, their principal focus is on character-recognition (e.g., Everson, 2011; Li, Gaffney, & Packard, 2002; McBride-Chang & Chen, 2003). While beginning readers of Chinese may have limited ability to process phrases and short sentences, assessing readers who have studied Chinese for longer may require examination of their reading ability at various levels. In other words, L2 Chinese reading may be better assessed via the inclusion of morpheme-discrimination, grammar, cloze, and passage-comprehension subtests, due to their high loadings in the SEM model of this study.

5.5 Strategy-Use in L2 Chinese Reading

Strategy use, considered as an aspect of higher-level processing, was measured in the current study using a survey developed by Phakiti (2007). The results support the study's hypothesis that strategy use can be divided into two types, cognitive and metacognitive, which were found to be highly correlated (0.71). The overlapping variance between these two constructs was approximately 50%, showing that it might be preferable to separate them, rather than clustering them into a single higher-level strategy-use construct. Moreover, the current

results revealed that metacognitive-strategy use, considered as a latent variable, did not directly regulate reading processes. Instead, strategy use was found to facilitate reading processes mainly through the latent variable of cognitive-strategy use.

The findings are in line with those of Phakiti (2008) and Purpura (1999), that a higher-order construct of cognitive-strategy use is directly regulated by another higher-order construct, metacognitive-strategy use, and that the latter indirectly facilitates reading comprehension through the former. When examining the nature of strategy use across two ESL assessment scenarios – a midterm and a final exam – Phakiti (2008) found the factor loadings of the SEM measurement models of cognitive-strategy use and metacognitive-strategy use varied to a small degree, with the cognitive model's factor loadings ranging from 0.81 to 0.87, and those of the metacognitive model only somewhat larger, at 0.69 to 0.89.

In comparison to Phakiti's (2008), the factor loadings of the present study's two measurement models were lower (0.41-0.80 for cognitive-strategy use and 0.46-0.74 for metacognitive-strategy use). One possible explanation is that, as Phakiti suggested, strategy use is volatile: being based on an individual's knowledge regarding the perceived purposes of the tasks at hand, and of how well certain strategies work together on certain tasks. In other words, in the current study, the low-stakes nature of the participants' tasks could have led to lower levels of perceived strategy use, in comparison to the relatively high-stakes midterm- and final-exam situations confronted by Phakiti's participants. However, the correlations between cognitive- and metacognitive-strategy use that Phakiti identified (approximately 0.75) was similar to the correlation between the same two constructs in the present study (0.71), and this may indicate that the executive function exerted by metacognitive strategies on cognitive strategies is stable across L2 reading-assessment contexts.

5.6 The Full SEM Model

Although many studies have investigated strategy-use in reading Chinese, the current study appears to be the first one to do so using SEM. The results of the initial SEM found non-significant correlations between cognitive-strategy use and comprehension skills, essentially meaning that the hypothesized SEM model functioned as two distinct systems in which connection between the two systems can be observed only by chance. However, this connection was strengthened after a path was added from monitoring strategy to character-recognition skills, on the basis of correlation-matrix and EFA results. From this point onward, the data supported the revised SEM model and all of its paths were statistically significant.

Taken as a whole, the results showed that 1) the three reading subtests significantly contributed to comprehension skills, with the passage-comprehension subtest exhibiting the highest factor loading; 2) the three character-recognition subtests significantly contributed to character-recognition skills, with the morpheme subtest exhibiting the highest factor loading; 3) all six strategies significantly contributed to either cognitive- or metacognitive-strategy use, with comprehending- and retrieval strategies exhibiting the highest factor loadings on cognitive-strategy use, and planning strategy exhibiting the highest factor loading on metacognitive-strategy-use; 4) a significant and positive relationship could be discerned between cognitive-strategy use and comprehension skills; 5) comprehension skills were significantly and directly explained by character-recognition skills and cognitive-strategy use; 6) another significant and negative relationship could be observed between monitoring strategy and character-recognition skills; and 7) taken together, character-recognition skills and the use of cognitive strategy, metacognitive strategy, and monitoring strategy explain 99% of the variance in the participants' comprehension skills.

The seemingly contradictory effects of monitoring strategy on character-recognition skills and of cognitive-strategy use on comprehension skills may be explained by Stavonich's (2000) interactive-compensatory model and the two functions of strategy-use. The interactive-compensatory model holds that any level of reading process can compensate for deficits in another level. It is possible that readers constantly monitor their comprehension status, and when comprehension deficits occur due to unfamiliar Chinese characters, metacognitive strategies are activated. Such strategies then allow the readers to assess their global comprehension and locate problematic areas (through evaluating strategy); select the best cognitive strategies and readjust steps for the employment of their mental resources (through planning strategy); and reach adequate comprehension (through comprehending, memory, and retrieval strategies). In sum, when monitoring strategies detect unfamiliar characters, other metacognitive strategies are activated for the purpose of regulating cognitive strategies, which in turn facilitate text-level comprehension, such as connecting details, reaching a coherent understanding of a text, and understanding the text's organization, in spite of deficits in character-recognition processes.

Moreover, this newly orchestrated text-level comprehension may allow readers to retrieve prior knowledge of the unfamiliar characters. A respondent's comment reported by Everson and Ke (1997) suggested the intricate interactions among the lower and higher levels of reading processes and strategy use:

> I'm trying to, when I come to a character I don't recognize, I'm saying it in my head, kinda mentally inserting a space while trying to keep the flow of the sentence going instead of just stopping at that one character, when just trying to help, seeing if that will trigger something in my memory that particular word, or character, that I'm not getting, which may be happening right now. (p. 8)

When considering the frequent interactions among reading processes, recognizing Chinese characters and words involves more mental resources than merely applying linguistic knowledge to grapheme decoding. To reach adequate reading comprehension, deciphered information from Chinese characters is necessary, but text-level comprehension can also help activate relevant vocabulary schema and retrieve the target vocabulary knowledge, which indicates that coherent reading comprehension is achieved through constant bi-directional processes between levels, rather than uni-directional bottom-up or top-down processes.

Often, negative correlations between strategy use and reading skills are ascribed to the compensatory function of strategy use, whereby strategies are employed to compensate for interrupted reading processes (e.g., Cohen, 2014). On the other side of the coin, positive correlations between the two are ascribed to the facilitative function of strategies, indicating that the higher one's perception of one's own strategy use is, the better one's reading comprehension will be (Afflerbach, Pearson, & Paris, 2008). Taking this line of thought further, the negative loading of monitoring → character recognition denotes the compensatory function of strategy use being triggered by interrupted character-recognition, and the positive loading of cognitive strategy use → reading comprehension indicates the facilitative function of strategy use toward better global comprehension of texts.

Jiang and Cohen (2012) reviewed studies of strategy use in L2 Chinese learning contexts, and identified four possible differences between strategy use in L2 Chinese reading and in L2 English reading. Among these four, two are pertinent to the current study: 1) the convergence of graphemes, phonemes, and meaning within one character, and 2) the relationships between sub-character components and morphemes. Before discussing the differences between the current study and Jiang and Cohen's, it should be noted that the strategies reviewed were mainly for

language learning across various contexts, and were not specific to language assessment. Although the current study did not support a path from cognitive-strategy use to character recognition, the measurement model of character-recognition skills indicated the participants' ability to systematically analyze sub-character structures for phonological and semantic information, as well as the importance of character-recognition processes to comprehension processes. In addition, based on verbal reports in Everson & Ke (1997), some introspective reports indicate strategy-use targeting at character-recognition, such as "Um, ok, it starts with, uh, 'a pair' … um this first word of main text, or the subtitle here, ah, looks like a measure. The '*fang*' something, I don't recognize that but I guessed at the pronunciation" (Everson & Ke, 1997, p. 7), "I'm trying to, when I come to a character I don't recognize, …" (Everson & Ke, 1997, p. 8), and "… this guy is the *waijiaobuzhang*, and, ok, so here we've got more information. So this the American Secretary of State or *guowu*?, or, I think, this American official, *Shierci*? (repeats), maybe it's *Scharci*, ah I don't' know, but I'm still sort of fixated on that name. But then I'll go down here." (Everson & Ke, 1997, p. 11), which show a combination of monitoring, evaluating, planning, retrieval, and memory strategies.

Based on the quoted verbal reports in Everson and Ke (1997), the reported processes appear to start with a combination of monitoring and evaluating strategies, so I examined the five monitoring and six evaluating strategies in the strategy-use survey and their descriptive statistics. Among the 11 strategies, Strategy 29 as a monitoring strategy, which is "I noticed when and where I was confused in the text," best reflects the initiation of the repairing processes, while the remaining 10 strategies are more general, such as Strategy 23 "I checked my own performance and progress as I moved along the test tasks" as an evaluating strategy. In addition, when the mean score of Strategy 29 is the highest (M = 3.28, SD = 0.64) among the 11 strategies (M =

2.91, *SD* = 0.49 for the other four monitoring strategies, and *M* = 2.78, *SD* = 0.48 for the six evaluating strategies), suggesting participants' most positive attitude to Strategy 29 and the path of Monitoring → character-recognition (see Appendix B for descriptive statistics of each strategy, and Strategy 29 is labeled as Monitor4 in Appendix B).

If the factor loadings of the latent variables in the revised SEM model are examined in light of the framework of communicative competence (Bachman & Palmer, 1996; Canale & Swain, 1980), it appears that comprehension skills and character-recognition skills belong to the domain of language knowledge, whereas strategy use belongs to the domain of strategic competence. This could further explain the relatively low contribution of strategy use to L2 Chinese reading-test performance, as compared to the contribution made by character-recognition skills.

Regarding the contribution of cognitive-strategy use to reading processing, the present project's findings paralleled those reported by some previous studies of L2 Chinese reading (Everson & Ke, 1997; Lee-Thompson, 2008) and L1 Chinese reading (Chan & Law, 2003). Like these L2 reading studies, the current study showed that both metacognitive and cognitive strategies were applied to lower- and higher-level reading processes, and also found another, similar contribution of strategy use to comprehension: specifically, 0.21 for the direct effect of cognitive-strategy use, 0.15 for the indirect effect of metacognitive-strategy use, and -0.33 for the indirect effect of monitoring strategy. On the L1 side, when investigating the relationship between metacognition and Chinese reading comprehension among L1 Cantonese children in Hong Kong, Chan and Law (2003) found significant and positive correlations between strategy use and comprehension (approximately 0.20).

The current study had similar findings to two important strategy-use studies conducted in L2 English settings (Phakiti, 2008; Purpura, 1999), in particular 1) a strong correlation between lower-level character-recognition processes and higher-level comprehension processes, 2) a direct effect of cognitive-strategy use on reading processing, 3) a strong correlation between the use of metacognitive and cognitive strategies, and 4) an indirect effect of metacognitive-strategy use on reading processing.

Unlike Purpura (1999), Phakiti (2008) measured lexico-grammatical reading ability via fill-in-the-blank cloze tests, and found that cognitive-strategy use directly contributed to performance on such tests, with contributions of 0.40 and 0.55 in the two settings studied. In the current study, the size of the direct effect of cognitive-strategy use on comprehension skills was 0.21. Although this is considerably smaller than that reported by Phakiti, the results of the current study nevertheless support the important role of strategy use in reading processing.

Although the current study identified a general function of strategy use that was similar to its general function identified in L2 English reading studies (Phakiti, 2008; Purpura, 1999), it also found a divergent specific function of monitoring. Purpura observed direct and positive effects of two metacognitive strategies, monitoring (0.10) and evaluating (0.47), on lower-level reading processing, whereas the current study found a direct and negative effect of monitoring strategy on lower-level reading processing (-0.34). This difference may be related to the participants having lower proficiency than Purpura's did, and/or to the differences between processing Chinese and English text.

5.7 Implications of the Present Study

The present study's interactive model of L2 Chinese reading has been supported, and it thus has a number of theoretical, methodological, and practical implications for L2-reading and language-testing research.

5.7.1 Theoretical implications. This study assumed that certain cognitive constructs of L2 Chinese reading are common to learners from various language backgrounds, and the acceptable model-fit indices of the revised SEM model support such an assumption. The study also simultaneously compared the contributions of character-recognition ability and strategy use to comprehension, and the results suggest that the importance of efficient character recognition overwhelms that of strategy use. However, mere comparison of these two contributions implies linear relationships between the two factors and comprehension, and overlooks the complicated relationships that exist between reading processes because strategy use interacts directly and indirectly with both character recognition and comprehension. In other words, when comparing the contributions of character recognition to those of strategy use, the nonlinear relationship of strategy use with the other processes is overlooked, and its supporting role in comprehension performance would be better captured via a nonlinear and interactive perspective on L2 Chinese reading.

In this study, strategy use was operationalized as conscious cognitive acts, and as having specific purposes in a particular context. As such, strategy use in the study differs from language-learning strategies, which are defined as "broad, teachable actions that learners choose from among alternatives and employ for L2 learning purposes" (Oxford, 2011, p. 12). Various scholars (Cohen, 2014; Phakiti, 2007; Tseng, Dörnyei, & Schmitt, 2006) have argued against the use of strategy-use surveys such as the Strategy Inventory for Language Learning (SILL)

(Oxford, 1990) in a wide range of learning contexts. Their objections all relate to the idea that, when used for a decontextualized purpose, what such surveys actually collect is the respondents' accumulated knowledge regarding strategy use, and they thus overlook the context-specific nature of online strategy-use as it occurs in people's working memory. To overcome the manifest ambiguities in existing definitions of language-learning strategies, and to emphasize learner autonomy, some scholars have begun switching to the notion of self-regulation (Oxford, 2011; Tseng, Dörnyei, & Schmitt, 2006); and Phakiti (2007) goes even further, arguing that the term "strategies" should be used only for online cognitive activities facilitating a particular L2 task performance. Following this line of reasoning, online strategy use for specific learning tasks might be conceptualized as part of strategic competence, and the meta-knowledge accumulated through online strategy use might further feed into self-regulatory competence for language learning.

In addition, scholars investigating communication strategies tend to adopt a view that strategy use is fundamentally a problem-solving technique (e.g., Cohen, 2014; Kasper & Kellerman, 1997; Macaro, 2001). This approach, by its nature, tends to foreground the compensatory function of strategy use and to downplay its facilitative functions. A theoretical implication of the negative and positive relationships between strategy use and reading processes identified by the present study could be uniting the compensatory and facilitative perspectives into a single framework that treats them as two sides of a coin. The compensatory function may be activated to *repair* insufficient comprehension or linguistic knowledge for the purpose of *facilitating* L2 task performance or linguistic competence. Therefore, both functions serve the same purposes, but with different connotations: the compensatory function focusing on the

distinction of deliberate strategies from automatic language skills; and the facilitative function emphasizing the positive contribution of strategies to L2 performance.

The present study drew on L1, L2, and HL reading development, L2 strategic competence, and L2 assessment research to investigate the relationships between selected cognitive constructs of L2 Chinese reading. In so doing, it first investigated the variance explained by language background (considered as a variable of higher-level reading processing) in L2 Chinese reading ability, character-recognition ability, and strategy use. An important theoretical implication of the results is that higher-level processing skills could be measured more accurately by tapping into the effect of language background on reading processing. On the other hand, although lower-level processing skills may be more language-specific, such skills can be developed quickly even by readers whose L1 orthography is very different from Chinese, and this may lead to lower-than-expected effects of language background on lower-level processing skills. Additionally, the significant group differences we observed in reading-test performance may have been linked to the nontechnical, narrative, and literary texts included in the reading test, which could have required greater activation of situational models of interpretation (Grabe, 2009) than the character-recognition test did. In other words, future investigators should bear in mind that large effects of language background on comprehension may be due to the activation of a situational model of reader interpretation, which is related to topical knowledge of the target culture more than to linguistic knowledge of the target language.

It should also be noted here that language background, as investigated in the present study, was operationalized based on Kondo-Brown's (2005) research, which suggested that, in terms of the amount of linguistic input, parents are significantly more beneficial for home-language development than grandparents. Based on those findings, the current dissertation

initially classified participants by language background based solely on their parents' native languages. However, this categorization scheme became problematic during screening of the responses to the language-background survey, in that Singaporean CMTLLs reported more Mandarin-use at home than did CFLLs. A reason for this might be that most Singaporean participants identified their parents as native speakers of English due to the prevailing language ideology in Singapore – to say nothing of the fact that, in a multilingual society like Singapore, the distinction among native language, first language, and dominant language can be less than clear. In addition, the language-background survey ignored the possibility that parents can be native speakers of more than one language. These issues may suggest that the status of being a foreign language learner or a heritage language learner might be better treated as being on a continuum for the target language exposure at home, or as an interval variable, rather than as categories or a nominal variable. Transferring language background categories to linguistic home exposure from different family members, i.e., multiple ordinal variables, to an interval variable could be achieved by generating factor scores through PCA and estimating person parameters through Rasch Analysis. In addition, the effect size of language background on different aspects of reading processing could be examined using multiple regression: by treating the factor scores of language background as the DV; the test scores of the reading subtests as IVs; and the R^2 of the multiple regression as the effect size of language background.

5.7.2 Methodological implications. This study operationalized reading acts as nonlinear and interactive, and combined multiple tasks that targeted various reading abilities, as proposed by Grabe and Jiang (2014). It also demonstrated the usefulness of SEM for investigating the nonlinear and interactive relationships between reading skills, especially in comparison to linear regression models and correlation comparisons. Specifically, when using multiple regression to

predict comprehension performance with the variables we examined, the regression formula would be comprehension performance = 0.99 * character-recognition + 0.21 * cognitive-strategy use + 0.15 * metacognitive-strategy use - 0.33 * monitoring strategy-use, which fails to describe the indirect and recursive interaction between strategy use and other reading processes.

5.7.3 Implications for practice. In addition to its theoretical and methodological implications, the present study has some practical implications for the assessment of L2 Chinese reading development. The most important of these relates to the findings regarding the factor loading between morpheme processing and character recognition, which complement those of Li and colleagues (2002) and Tong (2008): i.e., that morpheme-processing skills are likely to be a more sensitive measure of character-recognition skills, and thus of later reading comprehension, than are sub-character processing skills such as ortho-phonology and ortho-semantics. Researchers of Chinese reading development, whether L1 or L2, have always been interested in the level of statistical significance in the relation between sub-character processing and reading development. In this regard, the results of this dissertation have implications for test development, curriculum design, and instruction.

With regard to test development, it is suggested that in the case of lower-level comprehension skills, the most relevant instrument is the morpheme test, while the higher-level skills can best be assessed using a combination of grammar, cloze, and passage-comprehension tests. However, it would be useful to first investigate how morpheme processing varies across different test tasks, as well as how morpheme task effects interact with comprehension performance.

Because the results also indicated the existence of a compensatory relationship between character recognition and the use of a monitoring strategy, curriculum developers and teachers

should provide explicit instruction on morpheme processing and on monitoring one's affective and attentional state and comprehension level using a self-regulation approach (Oxford, 2011). For example, teachers might teach students how to monitor time, their state of concentration, and their level of understanding/confusion when they encounter unfamiliar Chinese characters, and also test them specifically on their grasp of these monitoring techniques. In this way, Chinese-language learners could potentially increase their morpheme-processing skill while simultaneously being exposed to a wide array of conscious acts that facilitate both lower- and higher-level reading. If this is indeed the case, exposure to online strategy use through explicit instruction would facilitate learners' knowledge of such strategy use as well as their strategic competence.

This dissertation found that the effect of language background was more clearly observable in the results of the morpheme, grammar, and cloze subtests. The administrators of Chinese-language programs that offer two tracks of curricula, one for HLLs and the other for FLLs, may wish to differentiate between these two groups as early as the placement-test stage. For such language programs, a reading instrument consisting of morpheme-discrimination, grammar, and cloze may be the most sensitive means of detecting the language-background effect. To simulate a situation involving a U.S.-based Chinese-language program, a follow-up profile analysis was conducted to investigate the language-background effect on these three subtests, using the three groups studied in this dissertation other than Singaporean CMTLLs. Group membership explained 20% of test-performance variance ($p < 0.001$), with Man-HLLs scoring the highest, Can-HLLs the second highest, and CFLLs the lowest on all three subtests. However, it is noteworthy that – even with this combination of reading tests – it was still challenging to differentiate among Can-HLLs, Man-HLLs and FLLs. This being the case, a

language-background survey tapping test takers' home Mandarin use might provide useful supplementary information for making placement decisions.

5.8 Limitations of the Study

This study has a number of limitations that should be taken into consideration when interpreting its findings.

5.8.1 Methodological limitations. This study's methodological limitations can be subdivided into concerns about participants, instruments and settings, and SEM analysis. First, the overall sample size was small, leading to even smaller sample sizes in each of the four subgroups, chiefly due to the specificity of the present study's requirements regarding language background. In addition, the participants were mainly recruited through personal connections, and were low-intermediate to low-advanced Chinese-language learners. As such, it remains to be seen if the results could be generalized to a larger and/or more randomly selected sample of learners with a wider range of abilities. Also, after screening the descriptive statistics, Singaporean CMTLLs emerged as an additional subgroup, introducing more confounding influence related to the variable of language background.

Second, in regard to measurement instruments, the findings may be specific to the use of multiple-choice reading tests, fill-in-the-blank cloze tests, and Likert-scaled questionnaires, as well as the low-stakes setting in which the study was carried out. Bachman and Palmer (1996) proposed six criteria by which to evaluate test validity, one of which was *authenticity*, i.e., the incorporation of the assessed individuals' language use in real-life situations. In this regard, multiple-choice test items are far from authentic, and a performance test targeting integrated language use would better reflect how L2 users actually use Chinese reading skills in real life. Although most of the prompts and items in the reading test and the character-recognition test

were validated for large-scale tests or research purposes, the respondents' test performances might have differed considerably if integrated performance tests had been administered instead (Adolf, Perfetti & Catts, 2011; Alderson, 2000). Moreover, when designing options of the multiple-choice items, I avoided items that could have multiple interpretations, and kept only those whose correct answers were clear and unambiguous. Authentic materials, in contrast, will tend to be subject to multiple interpretations, and avoiding such items in the reading test may have led to biased representation of target language use among test takers, given that generating multiple interpretations of texts is an aspect of the situational mode of reading comprehension.

Third, although the study aimed to simulate testing settings, online data collection was chosen in the hope of recruiting more participants, and this greatly reduced the control over the contexts in which the participants took the test. This in turn impacted on the reliability of the character-recognition test and strategy-use survey; therefore, the present study's findings may only be generalizable to test-preparation situations rather than actual testing.

Fourth, only part of the full range of reading skills was included in the present study. Other reading-processing skills involved in test-taking include parsing Chinese character strings into words, and predicting what Chinese characters will appear next based on preceding collocates and contextual cues. Moreover, though Oxford (2011) proposed that L2 learning involves use of cognitive strategies, affective strategies, and sociocultural-interactive strategies and their corresponding meta-strategies, only cognitive and metacognitive strategies were assessed in the dissertation. The assessment of online strategy use via a self-report questionnaire allowed it to be compared across all participants, but such questionnaires are less sensitive to and thus less able to capture the dynamic and complicated characteristics of such strategy use, as

compared to brain-scanning techniques, such as PET and fMRI scans, eye-tracking, and observations made by conversation analysis.

Lastly, the use of SEM analysis in this study was premised on the idea that common cognitive constructs are shared by participants from different language backgrounds. Although in the structural model of the revised SEM model, 99% of the comprehension-skills variance was explained by the 12 variables, significant χ^2 values indicated that the model significantly differed from the dataset, and that it was not the best fitting model to explain the data's underlying constructs. In this study, language background was regarded as an operationalized variable of background knowledge; and accordingly, the language-background variable was treated as a nesting variable. However, if background knowledge were deemed to be the extents of cultural and linguistic knowledge, it would be a crossed variable.

5.8.2 Conceptual limitations. Canale and Swain's (1980) concept of communicative competence minimally includes linguistic competence, strategic competence, and sociolinguistic competence. A major conceptual limitation of this dissertation is that sociolinguistic competence was not taken into consideration, and that cognitive variables played an arguably disproportionate role in the model. Testing scholars (Chapelle, 2012; Purpura, 2016) have suggested that testing pure linguistic competence would jeopardize test validity, because language use in real life involves resources other than linguistic competence. Therefore, it is important to develop a framework that can assess all three types of communicative competence, and which might include evaluation of sociolinguistic variables through a cognitive lens.

In the design of this study, as previously noted, it was assumed that there were common constructs of L2 Chinese reading across language backgrounds, and it was for this reason that all the participants were combined into one group for SEM analysis. The results of profile analyses

indicated that the effect of language background accounted for less than 20% of the variance in character-recognition ability, comprehension ability, or strategy-use in the study; and the model-fit indices of the revised SEM model suggested reasonable data-model fit. Nevertheless, it would probably be more effective to examine the common constructs of L2 Chinese reading through multi-group SEM analyses.

It should also be noted that all the participants in the current study were English-dominant, and L2 Chinese readers who are literate in Korean or Japanese may exhibit different constructs due to the closer relationships of their languages with Chinese. Even though the high correlations among the variables provide support for the hypothesized constructs of L2 Chinese reading, it is important that these findings be interpreted with a clear view of the current study's theoretical and practical contexts, i.e., 1) its cognitive and component perspective on reading ability, 2) its acceptance of a pedagogical view of reading as a language skill separate from listening, speaking, and writing, and 3) its focus on English-dominant bilingual Chinese readers who have received formal education in Mandarin.

5.9 Future Directions

Although this study used SEM analysis as a confirmatory tool for its constructs of L2 Chinese reading, it remained exploratory in nature, and its framework for examining component reading skills in the context of L2 Chinese reading should be treated as preliminary. In light of how much remains to be done to further the understanding of L2 Chinese reading, and the limitations of the current research, I would like to conclude this dissertation with some proposals for future avenues of research.

5.9.1 An interactive approach to reading in Chinese-language teaching. The results of the current study suggest that even in a low-stakes context, strategy use directly and indirectly

facilitates both comprehension and character-recognition skills. Far from arguing against the importance of character-recognition skills, this dissertation proposes that reading acts are both interactive and nonlinear, and recommends that online strategy use be clearly distinguished from the more context-free concept of language-learning strategies. Future work should seek to raise language teachers' awareness of the contextual basis and volatility of the strategy use that can be observed in every L2 task. In particular, future research on L2 Chinese reading will do a better job of capturing L2 Chinese readers' ability when it recognizes students as active and creative individuals, who are able 1) to consciously mediate lower- and higher-level reading processing, and 2) to adjust their cognitive processing based on the nature of their tasks and their learning purposes. Moreover, because the types of strategy use investigated in this study differ from the self-regulatory language-learning strategies proposed by Oxford (2011), future research could profitably investigate the relationship between context-specific online strategy use and context-free self-regulation competence, as doing so would improve the understanding of strategic competence in communicative language classrooms.

5.9.2 Authentic/real-life L2 Chinese reading. Under the framework of communicative language learning (Bachman & Palmer, 1996; Canale & Swain, 1980), it is important to link target-language use in real life to language tests. However, communicative competence is more easily related to oral communication skills and its relationship with productive literacy skill; and the relation of reading to communicative competence is less clear, especially for less-proficient language learners. It would be helpful to language teaching if we knew more about what less-proficient learners actually read, for what purposes, and how they use their reading comprehension to interact with others in real life, rather than viewing reading principally as a means to an end (i.e., the learning of a new language).

5.9.3 Exploring the effect of language background on test performance within a more social and communicative framework. This dissertation's results suggest that language background, which is related to socio-contextual factors, interacts with cognitive processing skills to varying degrees. Scholars of language testing (e.g., Chapelle, 2012; Kane, 2006; Shohamy, 2001) have argued strongly for more critically examining social influence on language assessment. One Man-HLL described by Wiley (2008) was born in Taiwan and raised in California, and refused to be posited as an incomplete native speaker of Mandarin in a placement test; indeed, he was so upset by this characterization that he said he might never return to Chinese language classes. Such examples clearly indicate that test validity could be endangered by social factors. One means of ensuring test validity in a specified social context would be adopting argumentative validity (Kane, 2006), and implementing a socially and contextually sensitive test is more likely to bring a positive washback effect, fostering the communicative competence necessary to overcoming the socio-contextual challenges faced by test users.

5.9.4 Treating language background as an interval variable. As mentioned above, PCA and Rasch Analysis could be used to transfer the multiple ordinal variables associated with home-language exposure from different family members to an interval variable; and the R^2 of a multiple regression could reveal the effect size of language background on different aspects of reading processing, if the factor scores of language background were treated as the DV and the reading-subtest scores as IVs.

5.10 Concluding Remarks

In conclusion, this study has attempted to examine constructs of L2 Chinese reading from an interactive perspective. Such constructs were broadly divided into lower-level and higher-level processing skills. The lower-level skills are related to character-recognition, and the higher-

level ones to comprehension skills, strategy use, and background knowledge operationalized as language background. The results demonstrated that the relationships between these skills are complex and nonlinear; and that some relationships are more easily observed than others, given the possible interactions that might occur. Moreover, while the identified contribution of strategy use to comprehension skills was relatively small, in comparison to its contribution to character-recognition skills, its effect was still significant: reflecting how readers actively engage a range cognitive resources to facilitate their comprehension. In spite of these complexities, the issues raised in this dissertation are of vital importance in the sphere of language learning and testing, if we wish to gain new insights on how human beings process and understand languages with a very different orthography from their own.

From a methodological perspective, this study has demonstrated the value of adopting an interactive view of L2 reading-ability assessment, and provides empirical findings regarding the relationships between the processing skills. It is hoped that the present study has contributed a new L2-acquisition perspective to the burgeoning field of L2 Chinese learning, and that this will allow for clearer comparisons of cognitive processing across various L2 contexts, and facilitate improvements in L2 learning and teaching.

REFERENCES

Adolf, S, Perfetti, C., & Catts, H. (2011). Developmental changes in reading comprehension: Implications for assessment and instruction. In S. Samuels & A. Farstrup (Eds.), *What research has to say about reading instruction* (4th ed., pp. 186–214). Newark, DE: International Reading Association.

Afflerbach, P., Pearson, P. D., & Paris, S. G. (2008). Between reading skills and reading strategies. *The Reading Teacher, 61*(5), 364–373.

Albirini, A., Benmamoun, E., & Saadah, E. (2011). Grammatical features of Egyptian and Palestinian Arabic heritage speakers' oral production. *Studies in Second Language Acquisition*, *33*, 273–303.

Alderson, J. C. (2000). *Assessing reading*. New York, NY: Cambridge.

Alexander, P. A, & Fox, E. (2013). Historical perspective on reading research and practice, redux. In D. E. Alvermann, N. J. Unrau, & R. B. Ruddell (Eds.), *Theoretical models and processes of reading* (pp. 3–46). Newark, DE: International Reading Association.

Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests*. Oxford: Oxford University Press.

Benmamoun, E., Montrul, S., & Polinsky, M. (2010). White paper: Prolegomena to heritage linguistics. Harvard University. Retrieved from Harvard University: http://scholar.harvard.edu/mpolinsky/publications/white-paper-prolegomena-heritage-linguistics

Bernhardt, E. B. (1991). *Reading development in a second language: Theoretical, empirical, and classroom perspectives*. Norwood, NJ: Ablex.

Block, D. (2003). What does the 'A' in SLA stand for? In *The social turn in second language acquisition* (pp. 92–120). Washington, D.C.: Georgetown University Press.

Bond, T., & Fox, C. M. (2015). *Applying the Rasch model* (3rd ed.). New York, NY: Routledge.

Bradley, D. (1992). Chinese as a pluricentric language. In M. Clyne (Ed.), *Pluricentric languages: Differing norms in different nations* (pp. 305–324). New York: Walter de Gruyter.

Brown, J. D. (2008). Effect size and eta squared. *Shiken: JALT Testing & Evaluation SIG Newsletter, 12*(2), 38–43. Retrieved from http://jalt.org/test/PDF/Brown28.pdf

Brown, J. D. (2009). Principal components analysis and exploratory factor analysis-Definitions, differences, and choices. *Shiken: JALT Testing & Evaluation SIG Newsletter, 13*(1), 26–30.

Byrne, B. (2010). *Structural equation modeling with AMOS: Basic concepts, applications, and programming* (2nd ed.). Hoboken, NJ: Taylor and Francis.

Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics, 1*(1), 1–47.

Chan, C. K. K., & Law, D. Y. K. (2003). Metacognitive beliefs and strategies in reading comprehension for Chinese children. In C. McBride-Chang & H. Chen (Eds.), *Reading development in Chinese children* (pp. 171–182). Westport, CT: Praeger.

Chao, Y. R. (1947). *Cantonese primer*. Cambridge, MA: Harvard University Press.

Chapelle, C. A. (2012). Conceptions of validity. In G. Fulcher & F. Davidson (Eds.), *The Routledge handbook of language testing* (pp. 21–33). New York, NY: Routledge.

Chen, M. J. (2003). Component skills for reading Chinese in primary school children. In C. McBride-Chang & H. Chen (Eds.), *Reading development in Chinese children* (pp. 19–33).

Westport, CT: Praeger.

Cheung, H., Chan, M., & Chong, K. (2007). Use of orthographic knowledge in reading by Chinese-English bi-scriptal children. *Language Learning, 57*(3), 469–505.

Cheung, H., & Ng, L. K. H. (2003). Chinese reading development in some major Chinese societies: An introduction. In C. McBride-Chang & H. Chen (Eds.), *Reading development in Chinese children* (pp. 3–17). Westport, CT: Praeger.

Chow, B. W.-Y. (2014). The differential roles of paired associate learning in Chinese and English word reading abilities in bilingual children. *Reading and Writing*, *27*(9), 1657–1672.

Cohen, A. D. (2014). *Strategies in Learning and Using a Second Language.* New York, NY: Routledge.

Cohen, A. D., & Upton, T. A. (2007). 'I want to go back to the text': Response strategies on the reading subtest of the new TOEFL. *Language Testing, 24*(2), 209–250.

College Board Advanced Placement Program. (2009). 2007 AP Chinese Language and culture released exam: Multiple choice questions. Retrieved April 9, 2015, from http://apcentral.collegeboard.com/apc/public/repository/2007_Chinese_Released_Exam_FINAL.pdf

Cook, V. J. (2002). Background to the L2 user. In *Portraits of the L2 user* (pp. 1–28). Buffalo, NY: Multilingual Matters.

Cook, V. J. (2015). Discussing the language and though of motion in second language speakers. *The Modern Language Journal, 99*(S1), 154–164.

DeFrancis, J. (1984). *The Chinese language: Fact and fantasy*. Honolulu: University of Hawai‘i Press.

Department of Statistics Singapore. (2015) *Population trends.* Retrieved from Department of Statistics Singapore website https://www.singstat.gov.sg/docs/default-source/default-document-library/publications/publications_and_papers/population_and_population_structure/population2015.pdf

Everson, M. (1998). Word recognition among learners of Chinese as a foreign language: Investigating the relationship between naming and knowing. *The Modern Language Journal, 82*(2), 194–204.

Everson, M. (2011). Literacy development in Chinese as a foreign language. In *Teaching Chinese as a foreign language: Theories and applications* (2nd ed., pp. 97–111). Boston, MA: Cheng & Tsui.

Everson, M., & Ke, C. (1997). An inquiry into the reading strategies of intermediate and advanced learners of Chinese as a foreign language. *Journal of the Chinese Language Teachers Association, 32*(1), 1–20.

Feldman, L., & Siok, W. (1999). Semantic radicals contribute to the visual identification of Chinese characters. *Journal of Memory and Language, 40*, 559–576.

Field, A. (2013). *Discovering statistics: IBM SPSS Statistics*. Los Angeles, CA: Sage.

Fleming, S., Hiple, D., & Ning, C. (2002). *Kan na! Authentic Chinese reading and video*. Retrieved from http://www.nflrc.org/login/scripts/material.php?id=56

Gorsuch, R. L. (1974). *Factor analysis*. Philadelphia: PA: W. B. Saunders.

Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. New York, NY: Cambridge University Press.

Grabe, W., & Jiang, X. (2014). Assessing reading. In A. Kunnan, & A. J. Kunnan (Ed.), *The Companion to Language Assessement* (1st ed., Vol. 1, pp. 185–200). New York, NY: John Wiley & Sons, Inc.

Hanban/Confucius Institute Headquarters. (n.d.) *Chinese Proficiency Test (HSK)*. Retrieved from Chinese Test website http://www.chinesetest.cn/gosign.do?id=1&lid=0

He, A. W. (2006). Toward an identity theory of the development of Chinese as a heritage language. *Heritage Language Journal*, *4*(1), 1–28.

He, A. W. (2008). Chinese as a heritage language: An introduction. In A. W. He, & Y. Xiao (Eds.), *Chinese as a heritage language: Fostering rooted world citizenry* (pp. 1–12). Honolulu: University of Hawai‘i, National Foreign Language Resource Center.

Hendryx, J. D. (2008). The Chinese heritage language learners' existing linguistic knowledge and abilities. In A. W. He, & Y. Xiao (Eds.), *Chinese as a heritage language: Fostering rooted world citizenry* (pp. 53–66). Honolulu: University of Hawai‘i, National Foreign Language Resource Center.

Ho, C. S.-H., Yau, P. W.-Y., & Au, A. (2003). Development of orthographic knowledge and its relationship with reading and speliing among Chinese kindergarten and primary school children. In C. McBride-Chang, & H.-C. Chen, *Reading development in Chinese children* (pp. 51–71). westport, CT: Praeger.

Hu, C. F. (2013). Predictors of reading in children with Chinese as a first language: a developmental and cross-linguistic perspective. *Reading and Writing*, *26*(2), 163–187.

Hudson, T. (2007). *Teaching second language reading*. Oxford: Oxford University Press.

Jiang, X., & Cohen, A. D. (2012). A critical review of research on strategies in learning Chinese as both a second and foreign language. *Studies in Second Language Learning and Teaching, 2*(1), 9–43.

Kagan, O., & Dillon, K. (2008). Issues in heritage language learning in the United States. *Encyclopedia of language and education*, *4*, 143–156.

Kane, M. (2006). Validation. In R. Brennan (ed.), *Educational measurement* (4th ed., pp. 17–64). Westport, CT: American Council on Education and Praeger.

Kasper, G., & Kellerman, E. (1997). Introduction: Approaches to communication strategies. In G. Kasper & E. Kellerman (Eds.), *Communication strategies: Psycholinguistic and sociolinguistic perspectives* (pp. 1–14). New York, NY: Longman.

Ke, C. (1998). Effects of language background on the learning of Chinese characters among foreign language students. *Foreign Language Annals, 31*(1), 91–100.

Kelleher, A. M. (2008). Placements and re-positionings: Tensions around CHL learning in a university Mandarin program. In A. W. He, & Y. Xiao (Eds.), *Chinese as a heritage language: Fostering rooted world citizenry* (pp. 239–258). Honolulu: University of Hawai‘i, National Foreign Language Resource Center.

Khalifa, H., & Weir, C. J. (2009). *Examining reading: Research and practice in assessing second language reading*. Cambridge, UK: Cambridge University.

Kline, P. (1994). *An easy guide to factor analysis*. New York, NY: Routledge.

Kline, R. B. (2011). *Principles and practice of structural equation modelling* (3rd ed.). New York, NY: Guilford.

Koda, K. (2002). Writing systems and learning to read in a second language. In W. L. Li, J. S. Gaffney, & J. L. Packard (Eds.), *Chinese children's reading acquisition* (pp. 225–248). Dordrecht, Netherlands: Kluwer Academic.

Koda, K. (2005). Learning to read across writing systems: Transfer, metalinguistic awareness, and second-language reading development. In V. Cook & B. Bassetti (Eds.), *Second language writing system*s (pp. 311–334). Tonawanda, NY: Multilingual Masters.

Koda, K., Lü, C., & Zhang, Y. (2008). Effects of print input on morphological awareness among Chinese heritage language learners. In A. W. He & Y. Xiao (Eds.), *Chinese as a heritage language: Fostering rooted world citizenry* (pp. 125–135). Honolulu: University of Hawai'i, National Foreign Language Resource Center.

Koda, K., Zhang. Y, & Yang, C.-L. (2008). Literacy development in Chinese as a heritage language. In A. W. He & Y. Xiao (Eds.), *Chinese as a heritage language: Fostering rooted world citizenry* (pp. 137–149). Honolulu: University of Hawai'i, National Foreign Language Resource Center.

Kondo-Brown, K. (2005). Differences in langauge skills: Heritage language learner subgroups and foreign language learners. *The Modern Language Journal, 89*(4), 563–581.

Ku, Y.-M. (2000). *Development of morphological awareness in Chinese and English* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses. (Order No. 3017133)

Kuo, L., & Anderson, R. (2008). Conceptual and methodological issues in comparing metalinguistic awareness across languages. In K. Koda & A. M. Zehler (Eds.), *Learning to read across languages* (pp. 39–67). New York: Routledge.

Lee, M. (1996). Non-Chinese heritage learners: Practices and implications. In X. Wang (Ed.), *A case study of Chinese heritage community language schools in the United States* (pp. 27–

31). Washington, DC: National Foreign Language Center.

Lee-Thompson, L.-C. (2008). An investigation of reading strategies applied by American learners of Chinese as a foreign language. *Foreign Language Annals, 41*(4), 702–721.

Li, D., & Duff, P. (2008). Issues in Chinese language education and research at the postsecondary level. In He, A. W., & Xiao, Y. (Eds.), *Chinese as a heritage language: Fostering rooted world citizenry* (pp. 13–32). Honolulu, HI: University of Hawai'i, National Foreign Language Resource Center.

Li, W., Anderson, R. C., Nagy, W., & Zhang, H. (2002). Facets of metalinguistic awareness that contribute to Chinese literacy. In W. Li, J. S. Gaffney, & J. L. Packard (Eds.), *Chinese children's reading acquisition: Theoretical and pedagogical issues* (pp. 87–111). Norwell, MA: Kluwer Academic.

Li, W., Gaffney, J. S., & Packard, J. L. (2002). *Chinese children's reading acquisition: Theoretical and pedagogical issues*. Norwell, MA: Kluwer Academic.

Li, Y. (1998). *Teaching and learning L2 reading: The Chinese case* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses. (Order No. 9900221)

Linacre, J. M. (2013). Facets computer program for many-facet Rasch measurement, version 3.71.0. Beaverton, OR: Winsteps.com

Liu, N. (2013). The linguistic vitality of Chinese in the United States. *Heritage Language Journal*, *10*(3), 294–304.

Lü, C., & Koda, K. (2011). The impact of home language and literacy support on English Chinese biliteracy acquisition among Chinese heritage language learners. *Heritage Language Journal*, *8*(2), 44–80.

Lynch, A. (2003). The relationship between second and heritage language acquisition: Notes on research and theory building. *Heritage Language Journal, 2*(1), 1–18.

Macaro, E. (2001). *Learning strategies in foreign and second language classrooms*. New York, NY: Continuum.

McBride-Chang, C., & Chen, H.-C. (Eds.). (2003). *Reading development in Chinese children*. Westport, CT: Praeger.

McBride-Chang, C., & Ho, C. (2000). Developmental issues in Chinese children's character acquisition. Journal of Educational Psychology, 92(1), 50-55.

McBride-Chang, C., & Kail, R. (2002). Cross-cultural similarities in the predictors of reading acquisition. *Child Development, 73*(5), 1392–1407.

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

McGinnis, S. (1996). Teaching Chinese to the Chinese: The development of an assessment and instructional model. In J. E. Liskin-Gasparro (Ed.), *Patterns and policies: The changing demographics of foreign language instruction* (pp. 107–121). Boston, MA: Heinle & Heinle.

Ministry of Education, Singapore. (2016a) *English Language and Literature.* Retrieved from Ministry of Education, Singapore website https://www.moe.gov.sg/docs/default-source/document/education/syllabuses/english-language-and-literature/files/english-primary-secondary-express-normal-academic.pdf

Ministry of Education, Singapore. (2016b) *Mother Tongue Language Policy.* Retrieved from Ministry of Education, Singapore website https://www.moe.gov.sg/admissions/returning-singaporeans/singaporeans-returning-home/mother-tongue-policy

Montrul, S. (2010). Current issues in heritage language acquisition. *Annual Review of Applied Linguistics*, *30*, 3–23.

Montrul, S. (2012). Is the heritage language like a second language? *EUROSLA Yearbook*, *12*, 1–29.

Montrul, S. (2013). How "native" are native speakers? *Heritage Language Journal, 10*(2), 15–39.

Morais, J. (1999). Constraints on the development of phonemic awareness. In S. A. Brady, & D. P. Shankweiler (Eds.), *Phonological processes in literacy: A tribute to Isabelle Y. Liberman* (pp. 5–28). Newark, DE: International Reading Association.

Mu, G. M. (2014). Learning Chinese as a heritage language in Australia and beyond: The role of capital. *Language and Education, 28*(5), 477–492.

Mueller, R. O., & Hancock, G. R. (2010). Structural Equation Modeling. In G. R. Hancock & R. O. Mueller (Eds.), *The Reviewer's Guide to Quantitative Methods in the Social Sciences* (pp. 371–384). Hillsdale, NJ: Lawrence Erlbaum Associates.

Nagy, W. E., Kuo-Kealoha, A., Wu, X., Li, W., Anderson, R. C., & Chen, X. (2002). The role of morphological awareness in learning to read Chinese. In W. L. Li, J. S. Gaffney, & J. L. Packard (Eds.), *Chinese children's reading acquisition* (pp. 59–86). Dordrecht, Netherlands: Kluwer Academic.

Norman, J. (1988). *Chinese*. New York: Cambridge University Press.

Ockey, G. J. (2014). Exploratory factor analysis and structural equation modeling. In A. Kunnan, & A. J. Kunnan (Ed.), *The Companion to Language Assessement* (1st ed., Vol. 3, pp. 1224–1244). New York, NY: John Wiley & Sons, Inc.

O'Grady, W., Lee, O.-S., & Lee, J.-H. (2011). Practical and theoretical issues in the study of heritage language acquisition. *Heritage Language Journal*, *8*(3), 24–40.

O'Malley, J. M., & Chamot, A. (1990). *Learning strategies in second language acquisition.* New York: Cambridge University Press.

Ortega, L. (2013). SLA for the 21st century: Disciplinary progress, transdisciplinary relevance, and the bi/multilingual turn. *Language Learning, 63*(S1), 1–24.

Oxford, R. (1990). *Language learning strategies: What every teacher should know*. New York: Newbury House.

Oxford, R. L. (2011). *Teaching and Researching Language Learning Strategies.* Harlow, UK: Pearson Longman.

Perfetti, C. A. (1999). Comprehending written language: A blueprint of the reader. In C. M. Brown & P. Hogoort (Eds.) *The neurocognition of language* (pp. 167–208). Oxford University.

Perfetti, C. A., & Dunlap, S. (2008). Learning to read: General principles and writing system variations. In K. Koda & A. M. Zehler (Eds.), *Learning to read across languages: Cross-linguistic relationships in first- and second-language literacy development* (pp. 13–38). New York: Routledge.

Perfetti, C. A., & Tan, L. H. (1999). The constituency model of Chinese word identification. In J. Wang, A. W. Inhoff, & H.-C. Chen, *Reading Chines script: A cognitive analysis* (pp. 115–134). Mahwah, NJ: Erlbaum.

Phakiti, A. (2003). A closer look at the relationship of cognitive and metacognitive strategy use to EFL reading achievement test performance. *Language Testing, 20*(1), 26–56.

Phakiti, A. (2007). *Strategic competence and EFL reading test performance*. New York, NY: Peter Lang.

Phakiti, A. (2008). Construct validation of Bachman and Palmer's (1996) strategic competence model over time in EFL reading tests. *Language Testing, 25*(2), 237–272.

Purpura, J. (1996). *Modeling the relationship between test takers' reported cognitive and metacognitive strategy use and performance on language tests.* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses. (Order No. 9620753)

Purpura, J. (1997). An analysis of the relationships between test takers' cognitive and metacognitive strategy use and second language test performance. *Language Learning, 47*(2), 289–325.

Purpura, J. (1999). *Learner strategy use and performance on language tests: A structural equation modeling approach*. Cambridge: Cambridge University Press.

Purpura, J. (2016). Second and foreign language assessment. *The Modern Language Journal, 100*(S1), 190–208.

Rupp, A. A., Ferne, T., & Choi, H. (2006). How assessing reading with multiple-choice questions shapes the construct: A cognitive processing perspective. *Language Testing, 23*(4), 441–474.

Schoonen, R., Hulstijn, J., & Bossers, B. (1998). Metacognitive and language-specific knowledge in native and foreign language reading comprehension: An empirical study among Dutch students in Grades 6, 8, and 10. *Language Learning, 48*(1), 71–106.

Shen, H. H. (2010). Analysis of radical knowledge development among beginning CFL learners. In M. E. Everson, & H. H. Shen (Eds.), Research among learners of Chinese as a foreign language (Chinese language Teachers Association Monograph Series, Vol. 4). (pp. 45–65). Honolulu, HI: University of Hawai'i, national Foreign Language Recourse Center.

Shen, H. H. (2013). Chinese L2 literacy development: Cognitive characteristics, learning strategies, and pedagogical interventions. *Language and Linguitics, 7*(7), 371–387.

Shen, H. H., & Bear, D. R. (2000). Development of orthographic skills in Chinese chidlren. *Reading and Writing: An Interdisciplinary Journal, 13*, 197–236.

Shen, H. H., & Ke, C. (2007). Radical awareness and word acquisition among nonnative learners of Chinese. *The Modern Language Journal, 91*(1), 97–111.

Shohamy, E. (2001). Democratic assessment as an alternative. *Language Testing*, *18*(4), 373–391.

Shu, H., & Anderson, R. (1999). Learning to read Chinese: The development of metalinguistic awareness. In J. Wang, A. W. Inhoff, & H.-C. Chen (Eds.), Learning to read Chinese script: A cognitive analysis (pp. 1–18). Mahwah, NJ: Lawrence Erlbaum Associates.

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

Song, X. (2005). Language learning strategy use and English proficiency on Michigan English Language Assessment Battery. *Spaan Fellow Working Papers in Second or Foreign Language Assessment*, *3*, 1–26.

Stanovich, K. E. (2000). *Progress in universtanding reading: Scientific foundations and new frontiers.* New York, NY: The Guilford Press.

Steering Committee for the Test Of Proficiency-*Huayu*. (n.d.) *TOCFL Reading (Vol. 2).* Retrieved from Steering Committee for the Test Of Proficiency-*Huayu* website http://www.sc-top.org.tw/mocktest_e.php

Sung, K. (2011). Factors influencing Chinese language learners' strategy use. *International Journal of Multilingualism, 8*(2), 117–134.

Swender, E., Conrad, D., & Vicars, R. (2012). *ACTFL proficiency guidelines 2012*. Alexandria, VA: American Council for the Teaching of Foreign Languages.

Tabachnick, B., & Fidell, L. (2012). *Using multivariate statistics* (6th ed.). Upper Saddle River, NJ: Pearson Education.

Tan, L., & Perfetti, C. (1997). Visual Chinese character recognition: Does phonological information mediate access to meaning? *Journal of Memory and Language, 37*, 41–57.

Tong, X. (2008). *The development of Chinese word reading: Relations of sub-character processing, phonological awareness, morphological awareness, and orthographic knowledge to Chinese-English biscriptal reading.* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses. (Order No. 3363236)

Tong, X., & McBride-Chang, C. (2010). Chinese-English biscriptal reading: Cognitive component skills across orthographies. *Reading and Writing, 23*, 293–310.

Tseng, W.-T., Dörnyei, Z., & Schmitt, N. (2006). A new approach to assessing strategic learning: The case of self-regulation in vocabulary acquisition. *Applied Linguistics, 27*(1), 78–102.

Tzeng, O. J. L. (2002). Current issues in learning to read Chinese. In W. Li, J. S. Gaffney, & J. L. Packard (Eds.), *Chinese children's reading acquisition: Theoretical and pedagogical issues* (pp. 3–15). Norwell, MA: Kluwer Academic.

Urquhart, A., & Weir, C. (1998). The theory of reading. In *Reading in a second language:*

Process, product, and practice (pp. 84–99). London: Longman.

Valdés, G. (2000). The teaching of heritage languages: An introduction for Slavic-teaching professionals. In O. Kagan & B. Rifkin (Eds.), *The learning and teaching of Slavic languages and cultures* (pp. 375–403). Bloomington, IN: Slavica.

Valdés, G. (2005). Bilingualism, heritage language learners, and SLA research: Opportunities lost or seized? *The Modern Language Journal, 89*(3), 410–426.

Villa, D. J. (1996). Choosing a "standard" variety of Spanish for the instruction of native Spanish speakers in the US. *Foreign Language Annals*, *29*(2), 191–200.

Wang, J., & Leland, C. H. (2011). Beginning students' perceptions of effective activities for Chinese character recognition. *Reading in a Foreign Language, 23*(2), 208–224.

Wang, J., Spencer, K., & Xing, M. (2009). Metacognitive beliefs and strategies in learning Chinese as a foreign language. *System, 37*, 46–56.

Wiley, T. G. (2008). Dialect speakers as heritage language learners: A Chinese case study. In D. M. Brinton, O. Kagan, & S. Bauckus (Eds.), *Heritage language education: A new field emerging* (pp. 91–106). New York: Routledge.

Wong, K. F., & Xiao, Y. (2010). Diversity and difference: Identity issues of Chinese heritage language learners from dialect backgrounds. *Heritage Language Journal, 7*(2), 153–187.

Wong, W.-W. (2011). *The learning of Chinese orthography and its centrality in learning Chinese as a foreign language* (Unpublished doctoral dissertation). The University of Hong Kong.

Xiao, Y. (2006). Heritage learners in the Chinese language classroom: Home background. *Heritage Language Journal*, *4*(1), 47–56.

Xiao. Y. (2008). Home literacy environment in CHL development. In A. W. He & Y. Xiao (Eds.), *Chinese as a heritage language: Fostering rooted world citizenry* (pp. 151–166). Honolulu: University of Hawai‘i, National Foreign Language Resource Center.

Xu, Y., Chang, L.-Y., & Perfetti, C. A. (2014). The effect of radical-based grouping in character learning in Chinese as a foreign language. *The Modern Language Journal*, *98*(3), 773–793.

Ye, D.-M. (2008). 華語語音學：語音理論 *Huayu yuyinxue: yuyin lilun* [Mandarin phonology: Phonology theories]. Taiwan: 師大書苑 *Shida shuyuan*.

Ye, F. & Dong, G. (2011, May 5). 专业太”新”难创业 *Zhanye tai “xin” nan chuangye* [Too “new” to be entrepreneurs]. *People’s Daily Overseas Edition*. Retrieved from http://paper.people.com.cn/rmrbhwb/html/2011-05/05/content_811613.htm

Zhang, L. (2014). College Chinese heritage language learners’ implicit knowledge of compound sentences. *Heritage Language Journal*, *11*(1), 45–75.

Zhang, D., & Koda, K. (2011). Home literacy environment and word knowledge development: A study of young learners of Chinese as a heritage language. *Bilingual Research Journal: The Journal of the National Association for Bilingual Education, 34*(1), 4–18.

Zhao, A., Guo, Y., & Dynia, J. (2013). Foreign language reading anxiety: Chinese as a foreign language in the United States. *The Modern Language Journal, 97*(3), 764–778.

Appendix A

The Complete Version of the Test (including the language background survey, the reading test, the strategy use survey, and the Chinese character test)

Language Background and Second Language Reading Processing in Chinese Reading Test Performance

Participant's Agreement to Participate

My name is Wei-Li Hsu. I am a PhD candidate at the University of Hawai'i at Manoa (UH), in the Department of Second Language Studies. The purpose of this research project is to evaluate the influence of language background on processing Chinese scripts for English speakers learning Mandarin.

If you meet the four criteria, you are invited to participate in a research study.
1. You are NOT a native speaker of Chinese and your strongest language is English
2. Older than 16 years old
3. Able to read paragraph-length in Chinese (between CHN 102 and CHN 401 at university, novice-high to intermediate level based on ACTFL, or between 3 and 5 in AP Chinese)
4. BOTH of your parents are (a) native speakers of English, (b) native speakers of Mandarin, or (c) native speakers of Cantonese.

What activities will you do in the study and how long will the activities last? If you participate, the test will last for about 60 minutes. There are four parts of the test. The first part is a language background surveyis; the second part is a reading comprehension test; the third part is about reading strategy; and the fourth part is about metalinguistic knowledge of Chinese.

Benefits and Risks: There is no direct benefit of participating this study and I believe there is only minimal risk to you in participating in this project. One potential risk is the uneasy feeling when answering more difficult reading comprehension questions. When encountering challenging comprehension questions, participants are encouraged to try their best, but there is no penalty to skip those questions.

Confidentiality and Privacy: I will not ask your name. The only personal identifiable information that may be linked to you is your email address which is used to receive the gift card. Your email address will be stored in a separate file, and it will not be linked to your responses. When I report the results of my research, there will be no personal information that would identify you. If you would like a copy of my final report, please contact me at the number listed near the end of this consent form.

I will keep all information of the online test in a safe place. Only I will have access to the information. Other agencies that have legal permission have the right to review research records. The University of Hawaii Human Studies Program has the right to review research records for this study.

Voluntary Participation: Participation in this research project is voluntary. You are free to choose to participate or not to participate in this project; and you can withdraw your permission without any penalty. By clicking the button of "Submit" at the end of this test, you agree to participate this study.

Questions: If you have any questions about this project, please contact me at via e-mail (whsu@hawaii.edu).

If you have any questions about your rights in this project, you can contact the University of Hawaii, Human Studies Program, by phone at (808) 956-5007 or by e-mail at uhirb@hawaii.edu.

* Required

1. **Please provide your E-mail if you agree to participate.** *
The email address is for sending your Amazon.com or Starbucks gift card upon completion.

2. **Which gift card would you prefer?** *
Mark only one oval.

- Amazon.com
- Starbucks

By clicking the button "Submit" at the end of this test, you agree to participate this study.

Language Background Survey

Please remember that in order to participate this study, you need to meet ALL of the four criteria.
1. You are NOT a native speaker of Chinese and your strongest language is English
2. Older than 16 years old
3. Able to read paragraph-length in Chinese (have learned Chinese for more than one year at university-level, between novice and advanced levels, or between 3 and 5 in AP Chinese)
4. BOTH of YOUR PARENTS are (a) native speakers of English, (b) native speakers of Mandarin, or (c) native speakers of Cantonese.

3. **BOTH of your parents are native speakers of**
Mark only one oval.

- English
- Mandarin
- Cantonese

4. **Your birth year**
EX: 1991

5. **Your birth place and country**
EX: Guangzhou, China

6. **In what year did you move to the United States, Canada or other English-speaking country, if you were not born in any of these countries.**
If you were born in any of the English-speaking countries, please fill in "N/A."

7. **Your MANDARIN use at home**
Mark only one oval per row.

	Not at all	Sometimes	Often	Very Often	Always
Mother					
Father					
Siblings					
Maternal grandparents					
Paternal grandparents					
Other relatives					

8. **Please rate your MANDARIN proficiency in each category.**
Mark only one oval per row.

	No proficiency	Minimal	Developing	Proficient	Exemplary
Listening					
Speaking					
Reading					
Writing					

9. **How many years have you been in a Chinese-speaking country?**

10. **How many years have you been learning Mandarin in Chinese language school? For example, Chinese weekend school.**

11. **How many years have you been learning Mandarin in middle school and high school?**

12. **How many years have you been learning Mandarin in college?**

13. **If you had received other kinds of formal learning of Chinese, please specify.**

14. **What other foreign language(s) have you learned, other than Mandarin? Please put the duration of learning that language in parentheses.**
EX: Cantonese (2 years)

The Chinese Reading Test

Introduction of the Chinese reading test

There are three tasks in the reading test: 8 multiple-choice grammar questions, 12 fill-in blanks, and 10 multiple-choice reading comprehension questions from 5 passages.

40 minutes is expected to finish the three tasks.

The first task-Grammar

Please choose the best answer to complete each sentence so the sentence is structurally correct and appropriate.

15. **1. 这个颜色很适合你，你可以__这件上衣。(這個顏色很適合你，你可以__這件上衣。)**
Mark only one oval.
- 试穿穿(試穿穿)
- 穿穿看
- 试试穿(試試穿)
- 试试看穿(試試看穿)

16. **2. 春天要到了，最近天气一天__一天暖和。(春天要到了，最近天氣一天__一天暖和。)**
Mark only one oval.
- 越
- 很
- 更
- 比

17. **3. 我们快迟到了，你准备好了__？(我們快遲到了，你準備好了__？)**
Mark only one oval.

- 啊
- 吗(嗎)
- 啦
- 呢

18. **4. __能符合规定，__这个计划无法实行。(__能符合規定，__這個計劃無法實行。)**
Mark only one oval.

- 因为...所以(因為...所以)
- 除非...才
- 除非...否则(除非...否則)
- 非...不可

19. **5. 你这样____在帮他，____在害他。(你這樣____在幫他，____在害他。)**
Mark only one oval.

- 是...是
- 不是...而是
- 是...而是(是...而是)
- 就是...是

20. **6. 这个工作很复杂，花了我很长时间才办__。(這個工作很複雜，花了我很長時間才办__。)**
Mark only one oval.

- 的
- 好
- 下
- 出

21. **7. 我朋友都__台湾电影很感兴趣。(我朋友都__台灣電影很感興趣。)**
Mark only one oval.

- 对(對)
- 所
- 才
- 是

22. **8.可以借我那__梳子吗? (可以借我那__梳子嗎?)**
Mark only one oval.

- 张(張)
- 个(個)
- 把
- 条(條)

The second task-Cloze test

(Please fill in ONE character in each blank)

Cloze-1

(Please fill in ONE character in each blank)

(Simplified)

钱塘潮

小芳，你好!

昨天农历八月十八日，我们一家都到钱塘江的边上去看有名的大 (9) 了，人多得不得了， (10) 纸上说有三十万 (11) 左右。在江边上，你 (12) 我我推你地在看这个大潮的景象，好看极了！我想世上没有比这个更好看的。看的人当中，各国人 (13) 有，真有意思！我们玩得很愉快，下星期见!

阿珍 九月十五日

==========

(Traditional)

錢塘潮

小芳，你好!

昨天農曆八月十八日，我們一家都到錢塘江的邊上去看有名的大 (9) 了，人多得不得了， (10) 紙上說有三十萬 (11) 左右。在江邊上，你 (12) 我我推你地在看這個大潮的景象，好看極了！我想世上沒有比這個更好看的。看的人當中，各國人 (13) 有，真有意思！我們玩得很愉快，下星期見!

阿珍 九月十五日

23. **(9)**

_____

24. **(10)**

---

25. **(11)**

---

26. **(12)**

---

27. **(13)**

Cloze-2

(Please fill in ONE character in each blank.)

(Simplified)

高速公路

现在全世界大约80多个国家有高速公路 。高速公路一般能适应每小时120公里或 (14) 高的速度，其发展 (15) 况往往可以看出一个国家的交通及经 (16) 水平。高速公路既有优点也有缺点，优 (17) 是行车速度快，安 (18) 方便，可以减少铁 (19) 等方面的交通压力；缺点是对环境影 (20) 大、收费高 。

==========

(Traditional)

高速公路

現在全世界大約80多個國家有高速公路 。高速公路一般能適應每小時120公裡或 (14) 高的速度，其發展 (15) 況往往可以看出一個國家的交通及經 (16) 水平。高速公路既有優點也有缺點，優 (17) 是行車速度快，安 (18) 方便，可以減少鐵 (19) 等方面的交通壓力；缺點是對環境影 (20) 大、收費高 。

28. **(14)**

---

29. **(15)**

---

30. **(16)**

---

31. **(17)**

.....

32. **(18)**

.....

33. **(19)**

.....

34. **(20)**

.....

The third task-Passage comprehension test

Read the following texts carefully for comprehension. Each is followed by some questions. Select the best answer according to the text.

Essay 1

Read this e-mail

Send

To... 爸妈

Cc...

Subject: 最近好吗?

亲爱的爸爸妈妈:

你们好！最近你们身体好吗？工作忙不忙？

我们快要考试了，所以我现在特别忙。从星期一到星期五，每天上午都去学校上课。星期二和星期四的下午我还有工作，别的下午我通常去咖啡馆看书。今天是星期六，我有空儿，给你们写信。学校的老师和同学都很好。我现在有很多朋友。保罗是我的新朋友，他上学期才从丹麦来的。以前也学过一点儿中文。

我住的地方很不错：附近有很多商店和饭馆儿，离汽车站也很近，吃饭，坐车，买东西都非常方便。房东对我们很好，常常请我们一起 吃饭。房东太太做的中国菜好吃极了，我特别喜欢吃她做的东坡肉。 到这儿以後，我每天都有运动，所以我身体很好，也没有病过，请爸爸，妈妈放心。

现在已经八点二十了，我该出门跟朋友见面了。 请常常给我来信。祝你们健康！

彼得
二零一五年三月十号

Send

To... 爸媽

Cc...

Subject: 最近好嗎?

親愛的爸爸媽媽:

你們好! 最近你們身體好嗎? 工作忙不忙?

我們快要考試了，所以我現在特別忙。從星期一到星期五，每天上午都去學校上課。星期二和星期四的下午我還有工作，別的下午我通常去咖啡館看書。今天是星期六，我有空兒，給你們寫信。學校的老師和同學都很好。我現在有很多朋友。保羅是我的新朋友，他上學期才從丹麥來的。以前也學過一點兒中文。

我住的地方很不錯: 附近有很多商店和飯館兒，離汽車站也很近，吃飯，坐車，買東西都非常方便。房東對我們很好，常常請我們一起 吃飯。房東太太做的中國菜好吃極了，我特別喜歡吃她做的東坡肉。 到這兒以後，我每天都有運動，所以我身體很好，也沒有病過，請爸爸，媽媽放心。

現在已經八點二十了，我該出門跟朋友見面了。 請常常給我來信。祝你們健康!

彼得
二零一五年三月十號

35. **21.**

What does Peter usually do on Wednesday afternoon?

Mark only one oval.

- preparing for tests
- going to classes
- staying in coffee shops
- studying in the library

Essay 2

Read this news article.

音乐欣赏将成为中国大学生必修课

记者在采访教育部负责人时得知，两年内所有中国高等院校都将开设音乐欣赏课程，而且这门课程将在五年内成为美个大学生的必修课。

据了解，通过和欧美大学生的比较，教育部的专家发现中国学生学习的课程大部份都是和专业直接有关的，对自己专业之外的东西了解极少，特别是在艺术方面，常常缺乏最起码的知识。欧美大学常常要求学生选修和专业无关的课程。例如一个学化学的学生必须学习午门人文科学课程。这样培养出来的学生，就不是简单的技术人员，而是全面的人才。

专家们还认为，现在中国学生常常听的是一些格调不高的流行音乐，这会让他们的思想受到不良影响。因此必须想办法让他们多听高水平的好音乐。不同於以前的唱歌和器乐演奏的音乐课，教育部这次设计的课程以欣赏欧洲十九世纪古典音乐为主，也有少量中国传统音乐。

考试的要求很简单，一般就是听一段音乐，让学生写出音乐的名字和作曲家的名字。教育部负责人认为，实现这一计划最大的困难就是师资，真正懂音乐的老师很难找到。

音樂欣賞將成為中國大學生必修課

記者在採訪教育部負責人時得知，兩年內所有中國高等院校都將開設音樂欣賞課程，而且這門課程將在五年內成為美個大學生的必修課。

據了解，通過和歐美大學生的比較，教育部的專家發現中國學生學習的課程大部份都是和專業直接有關的，對自己專業之外的東西了解極少，特別是在藝術方面，常常缺乏最起碼的知識。歐美大學常常要求學生選修和專業無關的課程。例如一個學化學的學生必須學習午門人文科學課程。這樣培養出來的學生，就不是簡單的技術人員，而是全面的人才。

專家們還認為，現在中國學生常常聽的是一些格調不高的流行音樂，這會讓他們的思想受到不良影響。因此必須想辦法讓他們多聽高水平的好音樂。不同於以前的唱歌和器樂演奏的音樂課，教育部這次設計的課程以欣賞歐洲十九世紀古典音樂為主，也有少量中國傳統音樂。

考試的要求很簡單，一般就是聽一段音樂，讓學生寫出音樂的名字和作曲家的名字。教育部負責人認為，實現這一計劃最大的困難就是師資，真正懂音樂的老師很難找到。

36. **22.**

Which of the following best describes the overall goal of the plan discussed in the article?
Mark only one oval.

- to produce college graduates who have some basic background in music
- to promote performance art in colleges
- to promote a revival of interest in traditional Chinese music
- to increase the number of students training to become music teachers

37. **23.**
What can we learn from the article?
Mark only one oval.

- Listening to music can increase students' IQ scores.
- Basic knowledge in arts is desirable for science students.
- Basic knowledge in traditional Chinese music is necessary to maintain Chinese heritage.
- Teachers are important for modern China.

38. **24.**
According to the article, how do Chinese college students differ from college students in Europe and the United States?
Mark only one oval.

- Relatively few Chinese students major in the field of Humanities.
- Chinese students tend to spend more time studying.
- Chinese students tend to take fewer courses outside their major field of study.
- Chinese students are generally better qualified for technical occupations when they graduate.

Essay 3

Read this short story.

小时候，奶奶给我讲过这样一个故事：很久很久以前，没有天也没有地，整个宇宙是混混沌沌的一大团，好像一个大皮球。有一个孩子就睡在这个大球中间。他睡啊睡啊，一睡就是九万八千年。有一天，不知道是什麽原因，这个孩子忽然醒了。他睁开眼睛一看，四周都是黑漆漆的一片，什麽都看不见。他伸手摸来摸去，摸到了一把宝剑。他挥舞宝剑，把大球划破了。大球里边的那些比较轻的东西往上升，变成了天;比较重的东西往下降，变成了地。天和地就这样分开了。

他担心天和地会再合上，就站在天和地的中间，用头顶着天，用脚踩着地。天和地之间的距离越来越大，这个孩子也越长越高，越长越壮，变成了一位巨人。这位巨人就这样站着，好像一根大柱子。过了一万八千年，巨人觉得实在太累了，再也坚持不住了，於是他就倒了下去。只听到“轰”的一声，他的头发变成了树林，肩膀变成了高山，肚子变成了平原，汗水变成了大海，血液变成了湖泊，口水变成了河流，眼泪变成了雨水，嘴里呼出的气变成了风和云。这样才有了我们这个美丽的世界。

小時候，奶奶給我講過這樣一個故事：很久很久以前，沒有天也沒有地，整個宇宙是混混沌沌的一大團，好像一個大皮球。有一個孩子就睡在這個大球中間。他睡啊睡啊，一睡就是九萬八千年。有一天，不知道是什麼原因，這個孩子忽然醒了。他睜開眼睛一看，四周都是黑漆漆的一片，什麼都看不見。他伸手摸來摸去，摸到了一把寶劍。他揮舞寶劍，把大球劃破了。大球裡邊的那些比較輕的東西往上升，變成了天;比較重的東西往下降，變成了地。天和地就這樣分開了。

他擔心天和地會再合上，就站在天和地的中間，用頭頂著天，用腳踩著地。天和地之間的距離越來越大，這個孩子也越長越高，越長越壯，變成了一位巨人。這位巨人就這樣站著，好像一根大柱子。過了一萬八千年，巨人覺得實在太累了，再也堅持不住了，於是他就倒了下去。只聽到"轟"的一聲，他的頭髮變成了樹林，肩膀變成了高山，肚子變成了平原，汗水變成了大海，血液變成了湖泊，口水變成了河流，眼淚變成了雨水，嘴裡呼出的氣變成了風和雲。這樣才有了我們這個美麗的世界。

39. **25.**
The hero's collapse is caused by
Mark only one oval.

- being hurt by the implement he grabbed earlier
- being stuck by the ball mentioned in the story
- extreme weather conditions on Earth
- separating earth and the sky for too long

Essay 4

Read this news article.

近几年，一些海归发现自己在海外所学的专业比较新，回国後很难找到与自己专业有关系的工作。海归们回中国很难找到工作可能有两种原因。

第一， 现在中国想要吸引移动通讯、新能源、生物医药专业的海归回来创业。如果是学习这些专业的海归，比较容易找到工作。但是如果是其他专业的海归找工作就比较困难。

第二， 没办法放弃学习多年的专业。从英国留学回来的李铃学的是产品管理，她自从今年一月回国就开始找工作，但一直没办法找到与自己专业有关的工作，许多公司都不需要这种专业的毕业生。

目前中国国内行业与他国教育的脱节是暂时的现象。同时，空白即机遇。行业的发展有一个从无到有、从小到大的过程，学习新兴专业的海归，应学会抓住这种机遇，坚持自己的理想，放远眼光，同时从基础做起，选择和自己的专业相关的基础性工作，一步步实现自己的梦想。

近幾年，一些海歸發現自己在海外所學的專業比較新，回國後很難找到與自己專業有關係的工作。海歸們回中國很難找到工作可能有兩種原因。

第一， 現在中國想要吸引移動通訊、新能源、生物醫藥專業的海歸回來創業。如果是學習這些專業的海歸，比較容易找到工作。但是如果是其他專業的海歸找工作就比較困難。

第二， 沒辦法放棄學習多年的專業。從英國留學回來的李鈴學的是產品管理，她自從今年一月回國就開始找工作，但一直沒辦法找到與自己專業有關的工作，許多公司都不需要這種專業的畢業生。

目前中國國內行業與他國教育的脫節是暫時的現象。同時，空白即機遇。行業的發展有一個從無到有、從小到大的過程，學習新興專業的海歸，應學會抓住這種機遇，堅持自己的理想，放遠眼光，同時從基礎做起，選擇和自己的專業相關的基礎性工作，一步步實現自己的夢想。

40. **26.**

Which field is best for 海归（海歸）？

Mark only one oval.

- oceanography
- business management
- product management
- mobile communications

Essay 5

Read this article.

相信很少有人想到南极考察人员在南极生存的最大威胁是极昼。毕竟在大家的意识里，皑皑的冰川丶极度的寒冷和急缺的食物才是考察人员面临的最大挑战。但事实上，他们的最大挑战并不是这些，而是那里的极昼。

所谓极昼，就是太阳终日都出现在地平线上的一种自然现象，它一般只会出现在夏季和冬季。当南极出现极昼时，北极就是极夜，反之亦然。

一位科学考查专家说："每当出现极昼时，没有了黑暗，也就没有了日期，工作人员连续几十天都生活在金灿灿的阳光下，人的生物钟一下子就彻底紊乱了。"因为人们都习惯了在夜晚的黑暗中睡觉，一旦失去了黑暗，那四周皑皑白雪和灿烂阳光让人很难闭上眼睛，即便你能睡上几分钟，也犹如在煎熬中。因此在南极，遭受雪崩和意外伤害的人，远没有被极昼伤害的人多。为了度过极昼期，考察人员做过很多尝试，比如加厚帐篷以增强帐篷内的阴暗度，甚至还尝试过在冰川和积雪下穴居等，但结果都不理想。

相信很少有人想到南極考察人員在南極生存的最大威脅是極晝。畢竟在大家的意識裡，皚皚的冰川、極度的寒冷和急缺的食物才是考察人員面臨的最大挑戰。但事實上，他們的最大挑戰並不是這些，而是那裡的極晝。

所謂極晝，就是太陽終日都出現在地平線上的一種自然現像，它一般只會出現在夏季和冬季。當南極出現極晝時，北極就是極夜，反之亦然。

一位科學考查專家說："每當出現極晝時，沒有了黑暗，也就沒有了日期，工作人員連續幾十天都生活在金燦燦的陽光下，人的生物鐘一下子就徹底紊亂了。"因為人們都習慣了在夜晚的黑暗中睡覺，一旦失去了黑暗，那四周皚皚白雪和燦爛陽光讓人很難閉上眼睛，即便你能睡上幾分鐘，也猶如在煎熬中。因此在南極，遭受雪崩和意外傷害的人，遠沒有被極晝傷害的人多。為了度過極晝期，考察人員做過很多嘗試，比如加厚帳篷以增強帳篷內的陰暗度，甚至還嘗試過在冰川和積雪下穴居等，但結果都不理想。

41. **27.**

What can we know about 极昼(極晝)?

Mark only one oval.

- It only exists in Antarctica.
- The sun never sets at that time.
- It often causes avalanche.
- It makes Antarctic investigators energetic.

42. **28.**
Which is the Antarctic investigator's description about their lives in Antartica?
Mark only one oval.

- The white night interrupts their biological clock.
- The midnight sun is especially spectacular.
- They got used to the environment after a fewmonths.
- Glaciers, extreme coldness, and scare food make surviving even more challenging.

43. **29.**
According to the short article, which way have Antarctic investigators tried to endure the harsh environment?
Mark only one oval.

- raise the temperature inside the tents
- live in caves underneath glaciers and snow
- use bright-colored tents
- none of above

44. **30.**
Which would be the best title of this article?
Mark only one oval.

- 生物钟(生物鐘)
- 两极生存挑战(兩極生存挑戰)
- 南极科学考察(南極科學考察)
- 南极极昼(南極極晝)

Survey of Reading Strategy Use

A number of statements which people use to describe themselves when they were taking a test are given below. Please read each statement and indicate how you thought during the test. Choose (Strongly Agree), (Agree), (Disagree) or (Strongly Disagree).

45. **Please rate how often you thought like the descriptions DURING taking the reading test.**
Mark only one oval per row.

	Strongly Agree	Agree	Disagree	Strongly Disagree
1. I planned what to do before I began to complete this reading test.				
2. I made sure I clarified the goals of the reading test tasks.				
3. I considered essential steps needed to complete the reading test.				
4. I made sure I understood what had to be done and how to do it.				
5. I knew what to do if my intended plans did not work efficiently while completing this reading test.				
6. I jumped back and forth through the reading test before I actually started to complete it.				
7. I tried to understand the relationships between ideas in the text and the tasks.				
8. I tried to understand the content of the text and tasks without looking up every word.				
9. I thought what was going to happen next while I was reading the text.				
10. I analyzed what the author meant or tried to say in the text.				
11. I tried to interpret hidden ideas/meanings in the text.				
12. I translated the text, tasks or questions into English.				
13. I summarized the main information in the text.				
14. I related the information from the text or tasks to my prior knowledge or experience.				
15. I reread texts or tasks several times when I felt I did not understand them.				
16. I knew which information was more or less important.				
17. I identified or guessed meanings of unknown words using context clues.				
18. I applied my learned grammar rules while reading and completing the reading tasks.				
19. I guessed meanings of unknown words using my knowledge of some character parts.				
20. I was aware of the time I have spent in this test.				

21. I knew how much of the reading and test tasks remained to be done while taking the test.				
22. I checked if I understood the texts and reading tasks.				
23. I checked my own performance and progress as I moved along the test tasks.				
24. I knew when I lost concentration while completing this test.				
25. I evaluated my plans or goals of my reading tasks constantly.				
26. I knew when I should read or complete the test more quickly or carefully.				
27. I double-checked my reading comprehension or performance.				
28. I immediately corrected my misunderstanding or performance mistakes when found.				
29. I noticed when and where I was confused in the text.				
30. I knew when I felt worried, tense or unmotivated to complete this reading test.				

Test of Chinese Character Knowledge

The first task-Pronunciation knowledge of Chinese characters

In each question, there will be a fake Chinese character and three options. The fake Chinese character consists of two real radicals. If the artificial character is real, which option would be closest in pronunciation?

1.

46. **1.**

Mark only one oval.

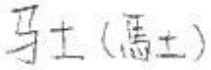
- 里
- 水
- 火

2.

47. **2.**
Mark only one oval.

- ◯ 木
- ◯ 成
- ◯ 本

3.



48. **3.**
Mark only one oval.

- ◯ 麦（麥）
- ◯ 马
- ◯ 土

4.

49. **4.**
Mark only one oval.

- ◯ 金
- ◯ 合
- ◯ 小

The second task-The meaning of Chinese radicals

In each question, there is a picture of an imaginary object. There are three options, each option uses an artificial Chinese character consisting of two real radicals. Please choose the option that best describes the object in the picture.

5. Please choose the option that best describes the imaginary object in the picture.

5.

A. 杈

B 枹

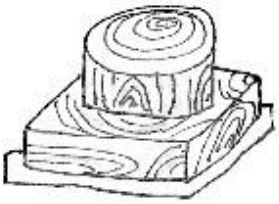
C 炊木

50. **5.**

Mark only one oval.

- A
- B
- C

6. Please choose the option that best describes the imaginary object in the picture.



6.

A. 木合

B 口木

C 合川

51. **6.**

Mark only one oval.

- A
- B
- C

7. Please choose the option that best describes the imaginary object in the picture.

7.

A 虫成

B 木虫

C 金木

52. **7.**

Mark only one oval.

- A
- B
- C

8. Please choose the option that best describes the imaginary object in the picture.

8.

A. 狪

B. 独

C. 蜩

53. **8.**

Mark only one oval.

- A
- B
- C

The third task-Character meanings

In the following 5 questions, there will be 3 phrases in each question. Each of these 3 phrases will contain one same character. Please choose the phrase which has the most different meaning from the other two.

For example, among (a)篮球/籃球, (b) 足球, and (c) 地球, the 球 in (c) has the most different meaning from the other two.

HINT Usually, a Chinese phrase, or a word in the sense of English, consists of two characters. The meaning of the character may remain the same when connected with other characters, or it may change. So please decide if the meaning of the character changes when it is used in the given phrase.

54. **9.**

Choose the one with most different meaning.

Mark only one oval.

- 冷气（冷氣）
- 和气（和氣）
- 空气（空氣）

55. **10.**
Choose the one with most different meaning.
Mark only one oval.

- 信心
- 信封
- 信箱

56. **11.**
Choose the one with most different meaning.
Mark only one oval.

- 校长（校長）
- 船长（船長）
- 生长（生長）

57. **12.**
Choose the one with most different meaning.
Mark only one oval.

- 可能
- 才能
- 功能

58. **13.**
Choose the one with most different meaning.
Mark only one oval.

- 口水
- 口红
- 口袋

Appendix B

Descriptive Statistics of Each Item in the Test

N = 85 Subtest	Possible Range	Min	Max	*Mean*	*SD*	Skewness *SE* = 0.26	Kurtosis *SE* = 0.52
Grammar1	0-1	0	1	0.48	0.50	0.07	-2.04
Grammar2	0-1	0	1	0.69	0.46	-0.86	-1.30
Grammar3	0-1	0	1	0.89	0.31	-2.61	4.92
Grammar4	0-1	0	1	0.62	0.49	-0.52	-1.77
Grammar5	0-1	0	1	0.77	0.43	-1.27	-0.40
Grammar6	0-1	0	1	0.66	0.48	-0.68	-1.57
Grammar7	0-1	0	1	0.75	0.43	-1.19	-0.59
Grammar8	0-1	0	1	0.58	0.50	-0.32	-1.95
Cloze1	0-1	0	1	0.27	0.45	1.05	-0.92
Cloze2	0-1	0	1	0.38	0.49	0.52	-1.77
Cloze3	0-1	0	1	0.55	0.50	-0.22	-2.00
Cloze4	0-1	0	1	0.44	0.50	0.27	-1.98
Cloze5	0-1	0	1	0.59	0.50	-0.37	-1.91
Cloze6	0-1	0	1	0.38	0.49	0.52	-1.77
Cloze7	0-1	0	1	0.46	0.50	0.17	-2.02
Cloze8	0-1	0	1	0.28	0.45	0.98	-1.06
Cloze9	0-1	0	1	0.53	0.50	-0.12	-2.03
Cloze10	0-1	0	1	0.51	0.50	-0.02	-2.05
Cloze11	0-1	0	1	0.28	0.45	0.98	-1.06
Cloze12	0-1	0	1	0.41	0.50	0.37	-1.91
Passage1	0-1	0	1	0.71	0.46	-0.92	-1.18
Passage2	0-1	0	1	0.53	0.50	-0.12	-2.03
Passage3	0-1	0	1	0.55	0.50	-0.22	-2.00
Passage4	0-1	0	1	0.67	0.47	-0.74	-1.49
Passage5	0-1	0	1	0.77	0.43	-1.27	-0.40
Passage6	0-1	0	1	0.41	0.50	0.37	-1.91
Passage7	0-1	0	1	0.62	0.49	-0.52	-1.77
Passage8	0-1	0	1	0.55	0.50	-0.22	-2.00
Passage9	0-1	0	1	0.55	0.50	-0.22	-2.00
Passage10	0-1	0	1	0.40	0.49	0.42	-1.87
Comp1	1-4	2	4	3.15	0.55	0.09	0.16
Comp2	1-4	2	4	3.36	0.61	-0.39	-0.63
Comp3	1-4	1	4	2.86	0.82	-0.26	-0.51
Comp4	1-4	2	4	3.02	0.64	-0.02	-0.46
Memory1	1-4	1	4	2.69	0.74	0.04	-0.40

Memory2	1-4	1	4	2.65	0.84	-0.36	-0.36
Memory3	1-4	1	4	2.79	0.66	-0.51	0.69
Memory4	1-4	2	4	3.36	0.57	-0.22	-0.71
Retrieval1	1-4	2	4	3.02	0.64	-0.30	0.53
Retrieval2	1-4	1	4	2.94	0.66	-0.19	0.05
Retrieval3	1-4	2	4	3.31	0.58	-0.52	1.65
Retrieval4	1-4	1	4	3.04	0.61	-0.34	1.02
Retrieval5	1-4	1	4	3.26	0.66	-0.59	0.53
Plan1	1-4	1	4	2.55	0.76	-0.18	-0.26
Plan2	1-4	2	4	3.02	0.49	0.06	1.40
Plan3	1-4	1	4	2.82	0.56	-0.87	1.97
Plan4	1-4	2	4	3.15	0.48	0.45	0.85
Plan5	1-4	1	4	2.76	0.63	-0.35	0.41
Plan6	1-4	1	4	2.74	1.00	-0.40	-0.86
Monitor1	1-4	1	4	2.86	0.79	-0.34	-0.21
Monitor2	1-4	1	4	2.55	0.78	-0.26	-0.28
Monitor3	1-4	2	4	3.18	0.58	-0.03	-0.20
Monitor4	1-4	2	4	3.32	0.60	-0.26	-0.60
Monitor5	1-4	1	4	3.15	0.61	-0.41	1.13
Eval1	1-4	1	4	2.86	0.56	-0.45	1.16
Eval2	1-4	1	4	2.72	0.70	-0.18	-0.04
Eval3	1-4	1	4	2.75	0.65	0.04	-0.27
Eval4	1-4	2	4	2.86	0.66	0.16	-0.67
Eval5	1-4	1	4	2.80	0.77	-0.12	-0.42
Eval6	1-4	1	4	2.95	0.62	-0.29	0.68
OrthoP1	0-1	0	1	0.78	0.42	-1.35	-0.18
OrthoP2	0-1	0	1	0.81	0.39	-1.62	0.65
OrthoP3	0-1	0	1	0.41	0.50	0.37	-1.91
OrthoP4	0-1	0	1	0.33	0.47	0.74	-1.49
OrthoS1	0-1	0	1	0.48	0.50	0.07	-2.04
OrthoS2	0-1	0	1	0.75	0.43	-1.19	-0.59
OrthoS3	0-1	0	1	0.54	0.50	-0.17	-2.02
OrthoS4	0-1	0	1	0.45	0.50	0.22	-2.00
Morpheme1	0-1	0	1	0.77	0.43	-1.27	-0.40
Morpheme2	0-1	0	1	0.75	0.43	-1.19	-0.59
Morpheme3	0-1	0	1	0.69	0.46	-0.86	-1.30
Morpheme4	0-1	0	1	0.46	0.50	0.17	-2.02
Morpheme5	0-1	0	1	0.59	0.50	-0.37	-1.91

Appendix C

Reliability of Each Subtest Using Rasch Analysis

	Cronbanch's Alpha			Rasch Analysis				
N = 85	*N*	*k*	Reliability	Anchored subtest	*n*	*k*[a]	*k*[b]	Reliability
Grammar	85	8	0.72	Passage	82	15	7	0.78
Cloze	85	12	0.91	NA	84	12	12	0.78
Passage	85	10	0.79	NA	80	9	9	0.70
Comprehending	85	4	0.65	Comprehending & Retrieval	67	4	4	0.68
Memory	85	4	0.31	Retrieval	81	8	4	0.66
Retrieval	85	5	0.44	Comprehending & Retrieval	77	5	5	0.68
Planning	85	6	0.52	Evaluating	84	11	6	0.77
Monitoring	85	5	0.60	Memory	72	9	5	0.72
Evaluating	85	6	0.75	Memory	72	10	6	0.80
Ortho-Phono	85	4	0.69	Passage	80	12	3	0.76
Ortho-Semantic	85	4	0.61	Passage	79	12	3	0.75
Morpheme	85	5	0.60	Passage	76	13	5	0.77

a: the original number of items in the subtest
b: the number of items after deleting misfitting items

Appendix D

Sample Sizes after Deleting Misfitting People

n	Total	Grammar	Cloze	Passage	Comp.	Memory	Retrieval	Plan.	Monitor.	Eval.	OrthoP	OrthoS	Morpheme
Singaporean CMTLLs	14	12	12	12	13	13	13	13	13	13	13	13	13
CFLLs	19	17	17	17	17	17	17	17	17	17	14	14	14
Man-HLLs	38	29	29	29	32	32	32	32	32	32	31	31	31
Can-HLLs	14	12	12	12	13	13	13	13	13	13	13	13	13
Subtotal	85	70	70	70	75	75	75	75	75	75	71	71	71

Appendix E

Descriptive Statistics with Rasch Ability Estimates

Table E1

Rasch Ability Estimates for the Reading Subtests

Reading Comp.	Grammar *M* (*SD*)	Cloze *M* (*SD*)	Passage *M* (*SD*)
Singaporean CMTLLs	1.01 (1.61)	-0.70 (2.08)	0.42 (1.79)
CFLLs	0.23 (2.10)	-0.65 (2.58)	0.70 (1.89)
Man-HLLs	0.43 (1.37)	-1.20 (2.84)	0.32 (1.91)
Can-HLLs	0.27 (1.45)	-1.94 (2.67)	0.31 (2.07)

Table E2

Rasch Ability Estimates for the Strategy-Use

Strategy Use	Comp. *M* (*SD*)	Memory *M* (*SD*)	Retrieval *M* (*SD*)	Eval. *M* (*SD*)	Monitor *M* (*SD*)	Plan *M* (*SD*)
Singaporean CMTLLs	0.14 (2.28)	0.12 (1.50)	0.01 (1.27)	-0.42 (1.29)	-0.14 (1.55)	-0.31 (2.11)
CFLLs	0.53 (2.15)	0.24 (1.62)	0.00 (1.67)	0.28 (1.67)	0.40 (1.20)	0.41 (2.15)
Man-HLLs	0.18 (1.87)	-0.20 (1.42)	0.03 (1.44)	0.12 (1.38)	0.16 (1.48)	0.14 (1.78)
Can-HLLs	0.40 (1.85)	0.42 (1.66)	0.31 (1.63)	0.08 (1.78)	-0.54 (1.72)	0.12 (2.12)

Table E3

Rasch Ability Estimates for the Character-Recognition Subtests

Character Recognition	Ortho-Phono *M* (*SD*)	Ortho-Semantic *M* (*SD*)	Morpheme *M* (*SD*)
Singaporean CMTLLs	0.06 (1.70)	-0.26 (1.85)	0.92 (1.16)
CFLLs	0.21 (2.55)	0.61 (1.88)	0.78 (1.63)
Man-HLLs	0.57 (2.00)	-0.06 (1.82)	0.19 (1.58)
Can-HLLs	0.62 (1.39)	-0.48 (1.52)	0.13 (2.12)