Comparing Reading Strategy Measures and L2 Readers' Performance on Different Comprehension Tasks

Yanjie Li [a]
Washington University in St. Louis
USA

Cindy Brantmeier [b]
Washington University in St. Louis
USA

Yanming Gao [c]
Northeast Normal University
China

Mark Hogrebe [a]
Washington University in St. Louis
USA

Abstract

In second language (L2) reading strategy research, two concerns need addressing: (1) the discrepancy in assessing strategy use between written surveys and verbal reports, and (2) the effect of using strategies on readers' comprehension outcomes when different types of comprehension tasks are utilized. The present study addressed these concerns by asking five more-proficient and five less-proficient Chinese English as a foreign language (EFL) university learners to read two domain-specific texts while assessing their strategy use. The results from both quantitative and qualitative analyses revealed a "gap" between the survey and verbal reports for assessing L2 reading strategy use. Additionally, the contribution of strategies to the readers' comprehension was found to vary by comprehension task type. Similarly, the variety of strategy use had a strong association with the readers' performance on only one of three types of comprehension tasks, free recall. The findings hold important implications on L2 reading strategy instruction.

Keywords: *L2 reading strategy, think-aloud protocol, different comprehension tasks, domain-specific texts, university-level Chinese EFL learners*

Strategy use is part of the reading process and has an indispensable impact on the readers' comprehension outcomes. Reading strategies are defined as behaviors, actions, and thoughts in which readers engage to augment their comprehension (He, 2008). Two types of measures, written surveys and verbal reports, have been commonly used for assessing strategy use in

second language (L2) reading research. Although both measures are widely used, some scholars have concerns about the discrepancy between the two measures for assessing strategy usage (Cohen, 2014; Endley, 2016). The present study attends to this concern by comparing the two measures for their differences and similarities in assessing strategy usage with university-level Chinese English as a foreign language (EFL) readers.

Brantmeier (2002) brought up another concern that much L2 reading research on reading strategies did not link readers' strategy use with their comprehension outcomes as indexed by readers' performance on assessment tasks (e.g., multiple-choice, free recall, summary test). This practice is, to date, not uncommon. Many studies on L2 reading strategy use either did not include any comprehension tests or had some tasks so that readers "read for general comprehension" but the results were not included in their data analyses (e.g., Alkhaleefah, 2017; Endley, 2016; Zhou & Zhao, 2014). In the present investigation, Chinese EFL learners' reading performance was assessed with three different comprehension tasks: free recall, sentence completion, and multiple-choice. The findings add new perspectives on the relationship between L2 readers' reading comprehension and strategy use.

Literature Review

Measures for L2 Reading Strategies

Both surveys and verbal reports are widely employed for assessing strategy use (e.g., Malcolm, 2009; Taki, 2016; Zhang & Wu, 2009 for using reading strategy surveys; Alkhaleefah, 2017; Endley, 2016; Lin & Yu, 2015 for using verbal reports). A survey of reading strategies consists of a number of statements asking participants to self-report their strategy use with a five-point Likert scale. For example, both Mokhtari and Sheorey's (2002) Survey of Reading Strategies (SORS) and Mokhtari and Reichard's (2002) Metacognitive Awareness of Reading Strategies Inventory (MARSI) include 30 statements asking readers to report their usage of global reading strategies, problem-solving strategies, and support reading strategies.

Verbal reports include the concurrent think-aloud protocol (TAP) and retrospective verbal reports, also called stimulated recall, conducted in the format of semi-structured interviews (Ericsson & Simon, 1993; Gass & Mackey, 2013). TAP asks readers to verbalize self-generated thoughts (Ericsson & Simon 1993). In other words, TAP asks what the readers would say to themselves in their head while performing the task. By comparison, stimulated recall is used after reading to prompt readers to recall thoughts they had while reading a text (Gass & Mackey, 2013). Some scholars are concerned that TAP procedures may distort participants' reading process. Dobrin's (1986) results do not support this concern. The Descriptive Test of Language Skills - Reading Comprehension Test was administered to readers under two conditions, standard operating conditions and in conjunction with verbal reports. No significant difference was found between the two administrations of the standardized test in participants' scores. TAP may not be a perfect procedure. Nevertheless, it reveals rich data about cognitive processes in a more direct manner which may not be achieved via other methods (Garner, 1987; Wade, 1990; Kong, 2006).

Both verbal reports and some reading strategy surveys (e.g., MARSI) have been validated and proved to be reliable measures (Brantmeier, 2002; Ericsson & Simon, 1993; Gass & Mackey, 2013; Mokhtari & Reichard, 2002). Nevertheless, these two types of measures have been observed to produce different findings on the same inquiry. For example, both Ghavamnia, Ketabi and Tavakoli (2013) and Zhang and Wu (2009) explored L2 strategy use, in terms of types and frequency, between the readers of different language proficiency levels. Using TAP and a semi-structured interview, Ghavamnia, Ketabi and Tavakoli (2013) concluded that, while the overall number of strategies used by proficient and less proficient readers was similar, the types of strategies used were different. In contrast, with a survey of reading strategies, Zhang and Wu (2009) found that the participants were similar in the overall types of strategy usage regardless of proficiency levels. Even though the conflicting findings might be related to one or more of the variables that were at play, the choice of reading strategy measurement is worthy of suspicion. The current investigation attempts to examine the possible "gap" between surveys and verbal reports for assessing L2 strategy usage.

Measures for L2 Reading Comprehension

In L2 reading research, some reading assessment formats include cloze, editing, free recall, matching activity, multiple-choice, ordering tasks, sentence completion, short answer, summary, and true/false (Alderson, 2000; Brantmeier, 2005, 2006; Grabe, 2009). The present study utilized three of these measures: free recall (FR), sentence completion (SC), and multiple-choice questions (MC).

FR asks readers to write down everything they can remember from the text without looking back at the passage. Since there are no retrieval cues to intervene between the reader and the text, FR is argued to be "a purer measure of comprehension" (Alderson, 2000, p. 230). In addition, it is generally agreed that the recall needs to be in the reader's dominant language, so the reading test does not become a writing test as well (Alderson, 2000; Bernhardt, 2011). To score FR, pausal units are commonly used (e.g., Liu & Brantmeier, 2019). A pausal unit is a unit that has a "pause on each end of it during normally paced oral reading" (Bernhardt, 1991, p. 208). FR represents a relatively subjective evaluation of the reader's comprehension. In comparison, MC is an objective method (Koda, 2005). MC requires readers to choose their answer to a question from four predetermined options, and there is no ambiguity in the scoring of the right/wrong answers (Brantmeier, 2006). A semi-objective alternative to MC is sentence completion, which asks participants to complete sentences according to the cues embedded in the sentence frames. The sentence frames must be worded so that all possible answers are foreseeable, and the objectivity of scoring depends upon the completeness of the answer key (Alderson 2000; Brantmeier, 2005, 2006).

The present study employed FR, SC, and MC to assess the participants' L2 reading comprehension because they are all commonly used, standardized assessment formats (Grabe, 2009). Besides, it is inadequate to use only one assessment method; rather, a variety of assessment tasks are necessary (Anderson, 2000; Brantmeier, 2005; Bernhardt, 1991). Alderson (2000) recommended that "objective methods can usefully be supplemented by more subjectively evaluated techniques" (p. 206). These three assessment formats constitute a well-balanced package of subjective, semi-objective, and objective methods.

Reading Strategies in L2 Reading

L2 reading is an extremely complex and interactive process during which readers capitalize on various available sources and exert a multitude of skills and strategies to construct meaning through interaction with written language (Bernhardt, 2011; He, 2008; Zhang & Wu, 2009). The critical role of reading strategies in the reading process has been stressed repeatedly. For instance, Cohen (2014) convincingly demonstrated that performance in language use depends in part on strategies. Further, Bernhardt (2011) suggested that reading comprehension strategies may be part of the remaining 50% of unexplained variance in her compensatory reading model.

Thus, many studies have examined the relationship between strategy usage and L2 reading while considering a number of variables, such as background knowledge (e.g., Bang & Zhao, 2007; Pritchard, 1990), first language (L1) literacy (e.g., Kong, 2006; Tsai et al., 2010), language proficiency (e.g., Carrell, 1989; Endley, 2016), linguistic distance (e.g., Block, 1986; Han & Stevenson, 2008), reading in L1 versus L2 (e.g., Lin & Yu, 2013; Yang, 2006), and reading outcomes (e.g., Anderson, 1991; Han, 2017). The focus of this paper is reading performance and its association with both the specific reading strategies that readers use and the variety and frequency of strategies used.

The use of L2 reading strategies has been studied extensively. Surprisingly, a majority of these studies did not include any measures of comprehension in their analyses (e.g., Akyel & Erçetin, 2009; Alkhaleefah, 2017; Chern, 1994; Endley, 2016; Ghavamnia et al., 2013; Kong, 2006; Lin & Yu, 2015; Malcolm, 2009; Mokhtari & Reichard, 2004; Yayli, 2010; Yang, 2006; Zhou & Zhao, 2014). In fact, there are only a handful of investigations concerning reading comprehension outcomes in examining strategy usage in L2 reading.

Among these studies, looking at strategy usage in terms of types and frequency, Anderson (1991) and Wang (2016) used multiple-choice to assess readers' L2 reading comprehension. Anderson's TAP data revealed that readers of different reading abilities reported using the same types of reading strategies. Echoing Anderson's results, Wang concluded that the most and least successful pairs of participants used a wide range of similar types of strategies with varying frequency of use in their reading and think-aloud processes. In both studies, multiple-choice questions were used to determine readers' comprehension, and neither study looked at the association between the use of specific strategies and reading performance.

It is exciting to see that two types of reading comprehension measures were included in a single study. In Tsai, Ernst, and Talley (2010), readers' L2 reading comprehension was measured with a cloze test and a multiple-choice test. Unfortunately, the study did not explore the direct link between the use of specific reading strategies and reading performance. Rather, the focus was on the differences in strategy use between skilled and less skilled readers, determined by the readers' language proficiency levels. Therefore, although they included two types of comprehension measures, Tsai, Ernst and Talley's analyses failed to bridge strategy use with L2 reading comprehension outcomes.

Research that includes at least three comprehension task types in a single study is scarce in this field. Brantmeier and Dragiyski (2009) used written recall, sentence completion, and multiple choice to investigate how the use of different reading strategies was associated with L2 reading

comprehension. Using MARSI, Brantmeier and Dragiyski showed that different combinations of reading strategies on the survey were positively associated with reading comprehension for specific assessment tasks. For example, the "try to picture and visualize information" strategy alone was found to be significantly associated with the reader's comprehension of the text only for multiple-choice; while the use of a combination of "try to picture and visualize information" and "try to guess the meaning of unknown words" strategies had a significant association with readers' performance on the recall measure only. Brantmeier and Dragiyski (2009) provided a long list of significant correlations between the use of a combination of different strategies and each reading comprehension measure. It seems apparent that the contribution of a specific strategy to reading comprehension largely depends on the specific comprehension measures employed. Brantmeier and Dragiyski contributed a great deal to our understanding of the relationship between the use of specific strategies and L2 reading comprehension. Nonetheless, as the study used a questionnaire for readers to self-report their strategy use, the results need to be interpreted with caution. Moreover, Brantmeier and Dragiyski did not inform us about the association between reading performance and strategy use in terms of frequency and variety.

The Present Study

Research Questions

The possible discrepancy between survey and verbal reports for assessing reading strategy use warrants studies comparing the two measures. Also, reading comprehension task type matters for understanding the dynamic association between strategy use and L2 reading performance. This mixed-methods study examines this concern and includes various reading comprehension measures while addressing the following research questions:

1. With Chinese EFL university learners reading two texts, is there a "gap" between a written survey and verbal reports (consisting of TAP procedures and semi-structured interviews) for assessing strategy usage?
2. With verbal reports, what is the relationship between reading strategies and comprehension, as measured with three types of assessment tasks: free recall, sentence completion, and multiple-choice?
3. With verbal reports, what is the relationship between the frequency and variety of strategy use and reading comprehension, as assessed with free recall, sentence completion, and multiple-choice?

Method

Participants

Second language reading imposes numerous unique challenges on learners. L2 reading is particularly arduous for students pursuing more specialized and professional knowledge at the university level as these readers are frequently required to read disciplinary texts (e.g., law, medicine, history) characterized as abstruse and conceptually dense (Koda, 2005). This

investigation sets out to help cultivate strategic readers and promote L2 learners' academic success in their field with medical students reading medical texts.

Verbal reports are qualitative in nature. L2 reading strategy research that involves verbal reports, therefore, tends to be composed of a relatively small number of participants, which are divided into groups of different proficiency levels to ensure variability. For example, Endley (2016) had 12 participants with five in the high proficiency reading group and seven in the low proficiency reading group, and Alkhaleefah (2017) had four participants consisting of two good readers and two poor readers. Likewise, the current study recruited ten participants from a large public medical university in Northeast China. Due to a small sample size, this study is limited in generalizability and therefore considered as a preliminary study, which can serve as a foundation for further research on this issue. To recruit the participants, we first chose 30 potential participants, who were asked to complete a demographic questionnaire in which participants reported their scores on the College English Test (CET), which is a national standardized test in China used to assess the English language proficiency of college students whose major is not English. Based on the ranking of their CET scores, ten undergraduates (two males and eight females, between 19 and 21 years old) from the two ends of the score range were selected to participate in the study. Each participant was rewarded 150 yuan (Chinese currency) for participating in the study.

Demographic Questionnaire

On the questionnaire, there were 17 questions about the participants' background information such as gender, age, and their scores on the CET test (see Appendix A) that were used to select the participants for this study.

Reading Texts

Brantmeier, Strube and Yu (2014) emphasized the importance of the text topic in the L2 reading process and proposed that studies that include domain-specific texts – texts that relate to a reader's area of study – may prove beneficial. The present study made an effort to examine how readers process their domain-specific texts in terms of strategy use by asking medical students to read two medical texts. Taken from the journal *Perspectives in Biology and Medicine*, the two texts used for this study were excerpts from two articles about Chinese traditional theories of drug interactions and the control of cardiovascular disease in the 20th Century, respectively (see Appendix B).

A professor who has taught pre-medicine courses for many years in a large Midwest private university in the United States recommended this journal because the readings conformed to those that medical students would be expected to comprehend immediately upon entering their studies. In addition, to illuminate the authentic reality of the reading process and comprehension of actual, authentic texts, the researcher chose not to modify or simplify the readings for this study. Moreover, based on students' self-report on the demographic questionnaire, only 25% of their disciplinary reading materials were in English. On average, the participants spent 2.6 hours weekly reading in English in general. The investigators hypothesized that the participants were

familiar with the topic but may be challenged by the language. This hypothesis was confirmed by the students in the interview.

Both texts had approximately 450 words and had a readability of 17, appropriate for college and graduate students, based on the Flesch Reading Ease Calculator.

Reading Comprehension Tasks

Following each text were three reading comprehension tasks: one free recall (FR), five sentence completion items (SC), and five multiple-choice questions (MC) (see Appendix B). FR asked participants to write as much as they could remember about each text in Chinese. Three specialists were consulted about the construction of the SC and MC items for measuring readers' comprehension of the text as a whole as well as the details (Brantmeier & Dragiyski, 2009).

Survey of Reading Strategies for L2 Contexts

The survey used in the current study adapted the SORS (Mokhtari & Sheorey, 2002) with minor modifications. SORS's reliability was tested with 147 English as a second language (ESL) students at a US college with an overall reliability of 0.89 (Mokhtari and Sheorey, 2008; Sheorey and Mokhtari, 2001, as cited in Malcolm, 2009). Three items in SORS were deleted because they were considered inappropriate for this study. First, the study asks participants to complete the survey based on the texts given. Two items, namely, "I have a purpose in mind when I read." and "I think about whether the content of the text fits my reading purpose." were therefore not relevant and not used. In addition, the TAP procedures were piloted with three students with similar backgrounds to the participants. All participants read the whole text out aloud during the TAP process. The item, "When text becomes difficult, I read aloud to help me remember it." as a strategy seemed not meaningful and thus was deleted as well. On the other hand, six strategies deemed necessary in the literature were added to the survey for the current study, that is, item 17: "I summarize what I read to reflect on important information in the text." (Mokhtari & Reichard, 2002), item 19: "When reading, I tend to identify main ideas and distinguish them from supporting ideas." (Lee-Thompson, 2008), item 30: "I tend to identify the text type and the structure or discourse patterns of the text." (Bakhshalinezhad et al., 2015), item 31: "When reading, I purposely ignore some unknown words and read on." (Hosenfeld, 1977), item32: "While reading, I constantly check if I know the main ideas of the text and clearly know it when there is a breakdown." (Block, 1986), and item 33: "I analyze grammatical structure (e.g., conjunctive adverbs and clauses) to help me understand the text." (Tsai et al., 2010). These added strategies did not overlap with those in SORS.

The final survey consisted of 33 items which were assigned into three groups: support strategies (9 items), problem-solving strategies (10 items), and global strategies (14 items) (see Appendix C). Participants chose a number from 1 ('I never or almost never do this') to 5 ('I always or almost always do this') on the survey using a 5-point Likert scale. The survey accompanied each of the two medical texts. Participants were clearly instructed to self-report the strategies they employed for comprehending the text they had just read.

Think-Aloud Protocols and Semi-Structured Interviews

TAP, a form of concurrent verbal reports, and semi-structured interviews, a form of retrospective verbal reports, served as the qualitative data source for analyzing reading strategies. In this study, both TAP procedures and semi-structured interviews were video recorded. The instructions for TAP followed Ericsson and Simon (1993), which asked participants to "TALK ALOUD while you are reading (…) say out aloud everything that you say to yourself silently. Just act as if you were alone in the room speaking to yourself." The participants were also instructed to point at what they were reading with the pen provided by the researcher. When the participant indicated they understood the instructions, the researcher asked the participant to demonstrate the TAP procedures as instructed with a short test whose topic and readability are similar to those used in the study. When indicating they were comfortable with the TAP procedures, participants proceeded with the reading task for the study.

Semi-structured interviews prompted participants to recall what was going on in the readers' minds at the moment of reading. The researcher carried out the interviews with each participant individually, as recommended by Gass and Mackey (2013). During the interviews, in addition to a few general questions about their experience of reading the text, the researcher asked participants to clarify some reading behaviors while playing back the recording. To give an example, one of the questions was "What were you thinking when you _____ (the reader's behavior)? Did you successfully _____ (the reader's intended goal) by _____ (the reader's behavior)?"

Procedure

All participants completed the demographic questionnaire on Day One. The participants were then scheduled to meet with the researcher individually two more times. On Day Two, each participant learned about TAP procedures and practiced with a short passage. When ready, the participant read Text One following TAP procedures and the process was video recorded. Note that to imitate students' real reading conditions when they read on their own, the researcher allowed the participants to use their cellphones as a dictionary. For the same reason, the researcher did not set a time limit for the reading task.

The researcher was present and sitting in the back of the room. The researcher did not intervene with participants' TAP process except for reminding the participant to 'Keep talking' when the participant stopped talking for approximately 15 seconds. Once finished reading, the participant signaled the researcher and proceeded to complete the three reading comprehension tasks following each passage. Directed by the instructions, the participant first recalled in Chinese both the main ideas and the details of the text. Then, cued by the sentence frames, the participant completed five sentences. Next, the participant selected the best answer to each multiple-choice question, five in total. Finally, after completing all three comprehension tasks, the participant was asked to self-report the extent to which they used each reading strategy on the survey for the text they had just read by circling a number from 1 ('I never or almost never do this') to 5 ('I always or almost always do this').

Meanwhile, the researcher managed the recording and took notes on the reading behaviors which needed clarification. In the end, the semi-structured interview was conducted and video recorded. On Day Three, the same processes were repeated for the second text.

Data Coding and Analyses

Both the TAP process and the semi-structured interviews were videotaped. Compared with audio recordings, the great advantage of video recordings is that both participants' verbalizations and their nonverbal reading behaviors (e.g., underlining and using a dictionary) were captured. The video content was first transcribed verbatim by the researcher. In addition to the researcher, a second scorer who had been teaching reading strategies to university-level Chinese EFL learners for decades also coded half of the data. The inter-rater reliability for Text One and Text Two was 95.49% and 97.48%, respectively. To aid the coding process, the raters used MAXQDA 2020, software for data coding and analysis. Both raters first segmented the transcripts based on idea units (Akyel & Erçetin, 2009). They then codified each segment using the reading strategy survey (see Appendix D).

Previous studies generally have audio recorded L2 readers' TAP procedures (e.g., Akyel & Erçetin, 2009; Endley, 2016; Lin & Yu, 2013; Yang, 2006; Yaylı, 2010). However, readers may use some strategies but not verbalize them. To capture participants' reading strategy use as precisely as possible, the researcher in this study 1) video recorded their TAP processes and 2) asked participants to specify whether they used certain reading strategies or not in the semi-structured interview. The video data revealed some strategies that were not readily caught on tape, such as "Glob2: I take an overview of the text to see what it is about before reading it." and "Glob3: I skim the text first by noting its characteristics like length and organization." In addition, during the interviews, participants identified their usage of certain strategies that were neither verbalized nor observed via video recordings, such as "Prob6: I try to picture or visualize information to help remember what I read." and "Prob8: When I read, I guess the meaning of unknown words or phrases."

Some strategies, such as those observed in the video recordings or identified in the interviews, were either used or not used by the participant. Other strategies were not appropriate to count. For example, all participants were observed re-reading. The content they re-read varied from a few words to a few sentences, and this pattern repeated itself every or every other sentence for the majority of the participants. For these reasons, this study's research team decided to adapt Lin and Yu's (2015) coding method and further grouped reading strategies into Yes/No strategies (19 out of 33) and Countable strategies (14 out of 33) (see Appendix D) for the purpose of data analysis. For the former, the code "1" was assigned to a strategy used by a participant and the code "0" was assigned if the strategy was not used; for the latter, the exact number of times the participant used the strategy was counted. Then, the following scores were calculated: (1) the number of times each countable strategy was used, (2) the percentage of strategy usage among participants for individual Yes/No strategies, (3) the number of unique reading strategies used for each participant (variety), and (4) the total number of strategies used for each participant (frequency).

As for reading comprehension measures, FR was scored based on the pausal unit protocol described in Bernhardt (2011). Two L1 English speakers read the two texts and marked 47 pausal units for Text One and 54 pausal units for Text Two. The pausal units independently marked by the two readers were identical for Text One and overlapped by 96% for Text Two, respectively. The disagreements were settled by a third reader. Each unit was worth one point.

Using retrieval cues, SC is an open-response task with certain limits placed on possible answers. All possible answers were determined by a L1 speaker of English and two experienced EFL teachers. MC asked participants to select one correct answer from four possible responses. The answers were pre-determined without ambiguity and the test-takers were not able to determine the correct responses by looking at the other questions on the page (Brantmeier, 2006). Each correct answer on SC and MC received one point.

Results

RQ1: With Chinese EFL university learners reading two texts, is there a "gap" between a written survey and verbal reports (consisting of TAP procedures and semi-structured interviews) for assessing strategy usage?

Countable Strategies

Table 1 presents the ranking of the 14 countable reading strategies and their corresponding mean scores reported on the survey (left column) and via verbal reports (right column). According to Mokhtari and Sheorey (2002), a mean score of below 2.5 on the survey is interpreted as low usage, 2.5 - 3.4 as medium, and 3.5 or greater as high. As Table 1 shows, among the 14 countable strategies, one strategy (Sup7) fell in the low usage category, three strategies (Glob6, Glob12, Sup5) fell in the medium usage category, and 10 strategies (Sup2, Prob2, Sup3, Sup1, Sup8, Sup4, Glob10, Glob11, Prob5, Glob8) fell in the high usage category. Through verbal reports, participants were observed using three strategies (Sup8, Sup3, Sup2) extensively and other strategies infrequently. The difference is striking. Six (43%) strategies (Glob12, Glob10, Sup5, Glob11, Sup4, Prob2) were reported as used at either the medium or the high level on the survey; by contrast, their observed usage was less than one time on average via verbal reports. This great difference demonstrates a clear discrepancy between the two measures for assessing the use of these countable strategies.

Table 1

Countable Reading Strategies Reported on the Survey Verses via Verbal Reports

Survey data			Verbal reports data		
Code	Identifier	Mean	Code	Identifier	Mean
Sup2	Underline	4.700	Sup8	Translate	206.00
Prob2	LoseConcent	4.450	Sup3	UseDictionary	102.60
Sup3	UseDictionary	4.300	Sup2	Underline	120.20
Sup1	TakeNotes	4.200	Prob5	Pause	9.20
Sup8	Translate	4.000	Glob8	EvaluateInfo	3.00
Sup4	Paraphrase	3.850	Sup1	TakeNotes	2.70
Glob10	Predict	3.550	Glob6	ContextClues	1.90
Glob11	CheckPrediction	3.550	Sup7	AskSelfQs	1.00
Prob5	Pause	3.550	Glob12	MainVsSupport	0.90
Glob8	EvaluateInfo	3.500	Glob10	Predict	0.20
Glob6	ContextClues	3.450	Sup5	Summarize	0.10
Glob12	MainVsSupport	3.400	Glob11	CheckPrediction	0.10
Sup5	Summarize	3.350	Sup4	Paraphrase	0.00
Sup7	AskSelfQs	2.050	Prob2	LoseConcent	0.00

Yes/No Strategies

Table 2 compares the percentage of participants who reported using certain Yes/No strategies. As shown in the survey column, except for two strategies (Glob5, Glob7) which were employed by 90% of the participants, all other strategies were said to be used 100% (at least occasionally). In contrast, with verbal reports, the use of strategies varied from 0 to 100%. A closer examination of the table shows that three strategies (Sup9, Prob4, Prob7) were reported as being used by 100% of the participants with both measures. On the other hand, for six of the Yes/No strategies (Glob2, Glob3, Glob9, Glob13, Glob14, Prob3), all participants reported employing them on the survey, but never used them according to verbal reports. The mismatch between the survey and verbal reports for assessing Yes/No reading strategies is clear.

Table 2

Percentage of Participants Who Reported Using the Yes/No Strategy on the Survey and via Verbal Reports

Code	Identifier	% on the survey	% via verbal reports
Sup6	Back&Forth	100	80
Sup9	ThinkL1&L2	100	100
Glob1	PriorKnowledge	100	50
Glob2	Overview	100	0
Glob3	SkimFirst	100	0
Glob4	What2Ignore	100	10
Glob5	UseTables	90	80
Glob7	TypographicalFeatures	90	80
Glob9	NewInfo	100	0
Glob13	DiscoursePtns	100	0
Glob14	FollowIdeas	100	0
Prob1	CarefulReading	100	60
Prob3	AdjustSpeed	100	0
Prob4	SlowWhenDiff	100	100
Prob6	VisualizeInfo	100	40
Prob7	Re-read	100	100
Prob8	GuessMeaning	100	30
Prob9	IgnoreUnknown	100	70
Prob10	AnalyzeGrammar	100	80

RQ2: With verbal reports, what is the relationship between reading strategies and comprehension, as measured with three types of assessment tasks: free recall, sentence completion, and multiple-choice?

Countable Strategies

Unlike null-hypothesis significance testing, effect sizes are not dependent upon sample size (Ferguson, 2009). As a measure of effect size, Pearson's r indicates the degree of shared variance between two variables and measures the strength of association between two variables. The interpretation of Pearson's r is context dependent, and in the contexts of social and behavioral research, Pearson's r of 0.2, 0.5, and 0.8 are considered as small, medium, large effect sizes, respectively (Ferguson, 2009; Lipsey et al., 2012). Correlations higher than 0.5 indicate that the two variables account for over 25 percent of shared variance (r^2), which is a meaningful amount of shared variance and warrants attention. Therefore, effect sizes above 0.5 were reported and discussed for this sample.

Table 3 presents the correlation results between strategy usage and the three reading measures (FR, SC, and MC). Four strategies, Sup1 ($r = 0.620$), Sup7 ($r = 0.665$), Glob10 ($r = 0.672$), and Glob11 ($r = 0.672$), had substantial contributions to the readers' performance on the FR measure. Two strategies, Sup7 ($r = 0.665$) and Sup8 ($r = 0.509$), contributed greatly to the MC measure, and one strategy, Sup3 (0.596), had a strong association with the readers' scores on the SC measure. In addition, strategy Prob5 was found to be negatively correlated with all comprehension measures: FR ($r = -0.683$), SC ($r = -0.760$), and MC ($r = -0.667$), suggesting that readers who paused more often performed worse on all three comprehension tasks. Given the small size of the study, the correlations are not meant to make a generalization to the population from the current sample. Rather, they are to be used as initial findings that demonstrate the importance of the research question and suggest further investigation with a larger sample.

Table 3

Correlations Between Reading Comprehension Scores (FR, SC, MC) and Individual Countable Reading Strategies

Code	Identifier	FR	SC	MC
Sup1	TakeNotes	**0.620**	0.364	0.433
Sup2	Underline	-0.160	0.061	0.398
Sup3	UseDictionary	0.344	**0.596**	0.405
Sup5	Summarize	-0.318	-0.396	-0.105
Sup7	AskSelfQs	**0.665**	0.379	**0.645**
Sup8	Translate	0.219	0.395	**0.509**
Glob6	ContextClues	0.263	0.206	0.119
Glob8	EvaluateInfo	0.480	0.291	-0.083
Glob10	Predict	**0.672**	0.452	0.419
Glob11	CheckPrediction	**0.672**	0.452	0.419
Glob12	MainVsSupport	0.001	-0.232	-0.434
Prob5	Pause	**-0.683**	**-0.760**	**-0.667**

Note. Pearson correlation was performed.
Effect sizes (r) > 0.5 are considered meaningful and bolded in the table.

Yes/No Strategies

Ten strategies did not contribute to the variability in reading scores because they were seldom used or not used at all. For the other nine strategies, independent sample t-tests were performed with the three reading measures for each strategy, comparing readers who used the strategy and those who did not use the strategy. As presented in Table 4, the results showed that participants who used Glob5 ($p = 0.003$) and Prob10 ($p = 0.003$) scored significantly higher on SC. Additionally, those who employed Glob7 ($p = 0.030$) scored significantly higher on MC. None of these Yes/No strategies had a significant contribution to FR.

Table 4

Significant Results of Independent Sample T-tests on the Three Reading Measures for Each Yes/No Strategy

Code	Identifier	Reading Measures	Mean Difference	Std. Error Difference	t	P
Glob5	UseTables	SC	3.500	0.802	4.365	0.003
Glob7	TypographicalFeatures	MC	3.250	1.234	2.633	0.030
Prob10	AnalyzeGrammar	SC	3.500	0.802	4.365	0.003

RQ3: With verbal reports, what is the relationship between the frequency and variety of strategy use and reading comprehension, as assessed with free recall, sentence completion, and multiple-choice?

Table 5 presents the results for Pearson's correlation between each reading measure and both the frequency and variety of strategy usage. No outstanding effect sizes were found for frequency. The association between variety and FR was notable ($r = 0.670$), suggesting that readers who performed better on FR used a greater variety of strategies.

Table 5

Correlations Between Different Reading Scores and the Frequency and Variety of Reading Strategies

	FR	SC	MC
Frequency	0.160	0.386	0.486
Variety	**0.670**	0.378	0.156

Note. Pearson correlation was performed.
Effect sizes (r) > 0.5 are considered meaningful and bolded in the table.

Discussion

RQ1: With Chinese EFL university learners reading two texts, is there a "gap" between a written survey and verbal reports (consisting of TAP procedures and semi-structured interviews) for assessing strategy usage?

The "gap" between the survey and verbal reports for assessing reading strategies reflects the differences in the nature of the two measures for assessing reading strategies. Reading strategy surveys ask participants to self-report their strategy usage. As "self-report is not grounded in specific behavior" (Cohen, 2014, p. 80), Baker and Brown (1984) have warned researchers not to "take students' word for what they do while reading" and pointed out that readers may often know that a strategy is effective, despite not using it themselves (cited in Carrell, 1989). The actual strategy usage is cued and impacted by numerous factors, such as text types (Grabe, 2009), readability of the texts (Mokhtari & Reichard, 2004), language proficiency (Brantmeier et al., 2014), reading purposes (Grabe & Stoller, 2011) and instructional experience (Taki, 2016).

Indeed, compared with the survey that asks readers to self-report their reading strategies, it is argued that verbal reports (e.g., TAP) more accurately reflect what readers actually do (Cohen, 2014). A few decades earlier, Bereiter and Bird (1985) had advocated for TAP and stated that 'thinking aloud offers the promise of breaking into the reading process to reveal on-line strategies' (p. 132, as cited in He, 2008). Recently, Durning et al.'s (2013) data collected with the Functional Magnetic Resonance Imaging (fMRI) technique adds evidence to support the notion that TAP is a reasonable measure of thought processes. Therefore, think-aloud methods are considered to be more effective compared to other methods involving introspection and retrospection (Ericsson and Simon, 1993). It stands to reason that strategy use indicated via verbal reports in this study more accurately reflects the reader's actual strategy use. Although the researcher instructed the participants to self-report their actual strategy usage with the survey based on the text they had just read, they perceived that they used more strategies than they actually did, as unveiled in verbal reports. The basis on which the participants made their decision on the number selected for each strategy on the survey is beyond the scope of this study. Future research is needed to explore participants' rationale behind their self-reported strategy use. This may shed light on the gap between the two reading strategy measures and on L2 reading process.

Despite the gap, the survey data suggest that participants recognized all 33 strategies in the survey. A similar finding was also reported by Alsheikh and Mokhtari (2011) and Malcolm (2009). The participants were aware of all the strategies; however, their strategy usage as measured with verbal reports was mostly limited to six reading strategies (Sup2: Underline, Sup3: UseDictionary, Sup8: Translate, Sup9: ThinkL1&L2, Prob4: SlowWhenDiff, Prob7: Re-read) for word- and sentence-level comprehension. This limited usage of strategies can be due to many reasons. Two of the possible explanations for this result were proposed by the data. First, the difficulty of the text limited the type of strategies the readers used (Yayli, 2010). No significant difference in reading comprehension scores was found between the participants of different proficiency levels, implying that the texts were too difficult for all participants and the discrepancy in reading performance was diminished. In addition, participants' average self-reported difficulty of the text in this study was 6.15 out of 10 (1 means "very easy"; 10 means

"don't understand at all") indicating the texts were relatively difficult for the participants. Grabe & Stoller (2011) warned that high linguistic demands may cause a text to be too difficult to read. Participants clearly conveyed the challenges associated with unfamiliar vocabulary and complex sentence structures during the interviews. The following extracts illustrate the problem:

> "I found it very difficult, because there may be many words in a sentence that I did not recognize, and I still could not translate the sentence after I looked up those words." (Participant A)

> "I did not understand many of the words. There were also some sentences that were very confusing." (Participant B)

This study coincides with Endley (2016) on the common reading problems that participants experience: problems with word recognition, difficulty in parsing complex grammatical structures, and failure to build words into higher-level meaning. As a result, participants were forced to pay substantial attention to tackling problems pertaining to the lower-level linguistic processes, including lexical access, syntactic parsing, and semantic proposition formation (Grabe & Stoller, 2011). Subsequently, as readers employ strategies to address reading problems or goals (Anderson, 2009), participants' strategy usage in this study was largely limited to a few strategies that targeted decoding problems.

Second, L2 strategy use is impacted by readers' socio-cultural background and instructional experience (Ghavamnia et al., 2013; Grabe & Stoller, 2011; Mokhtari & Reichard, 2004; Taki, 2016). In China, young Chinese students are regularly evaluated by numerous regional and national high stakes standardized exams for different purposes (Liu & Brantmeier, 2019). The washback effect seems to be reflected in strategy usage. The following extracts illustrate the issue:

> "Like when doing reading practice for the College English Tests 4 and 6, in my mind I predict this sentence is more important because I might be tested on the information later. For example, I circled 'not been' here due to this way of thinking." (Participant F)

> "I mark the words that I believe important. It is like doing the practicing reading tasks for the College English Test 6 test prep, and I would mark the common and frequent adjectives and verbs which are often seen in the test questions." (Participant I)

 As the English examinations in China commonly emphasize reading and writing abilities, vocabulary and grammar are valued greatly. Chinese students seldom read in English for fun; instead, they are commonly trained to read English texts carefully and thoroughly for the purpose of learning new vocabulary and grammar. Chinese students, therefore, tend to be dictionary-dependent and accuracy-oriented in reading (Chern, 1994). Not surprisingly, this reading practice familiarizes Chinese readers with decoding strategies, which may in part explain why participants favored the six bottom-up strategies in this study, regardless of proficiency levels.

RQ2: With verbal reports, what is the relationship between reading strategies and comprehension, as measured with three types of assessment tasks: free recall, sentence completion, and multiple-choice?

Several significant findings can be inferred from the results. Perhaps the most significant finding is that the association of reading strategies with reading comprehension varies depending upon the assessment techniques employed. This finding echoes Brantmeier and Dragiyski's (2009) conclusion that different strategies or combinations of strategies were positively associated with different reading comprehension measures. In the current study, our data show that reading strategies Sup1 (TakeNotes), Sup7 (AskSelfQs), Glob10 (Predict), and Glob11 (CheckPrediction) were related to FR scores; Sup3 (UseDictionary), Glob5 (UseTables), and Prob10 (AnalyzeGrammar) were related to SC scores; and Sup7 (AskSelfQs), Sup8 (Translate), and Glob7 (TypographicalFeatures) were related to MC scores. Some of the strategies would have been neglected if we had included only one of the reading measures in this study. Likewise, other types of comprehension measures such as cloze tests and summary tests might have foregrounded other reading strategies. Indeed, different measures captured different components or aspects of reading comprehension (Alderson, 2000; Grabe, 2009). Reading is a complex activity involving lower-level processes (e.g., lexical access) and higher-level processes (e.g., situation model of reader interpretation) (Grabe & Stoller, 2011). The various factors involved in L2 reading related to texts (e.g., text readability) and a reader's individual experiences (e.g., socio-cultural backgrounds) make the meaning-making process dynamic, recursive, and multifaceted (Grabe & Stoller, 2011; Koda, 2005; Tsai et al., 2010; Zhang & Wu, 2009). The challenges and goals for readers are varied when navigating texts. Therefore, the decisions pertaining to reading strategy usage and the ensuing effects on reading performance tend to be complex. As each individual assessment task provides a limited presentation of reading comprehension (Brantmeier, 2005, 2006), to explore the relationship between strategy usage and L2 reading performance comprehensively, a combination of different types of reading comprehension measures should be considered in a single study (Alderson, 2000; Bernhardt, 1991).

Second, readers pause more often and more likely suffer comprehension failure when their linguistic knowledge is limited. The notably negative correlation between strategy Prob5 (Pause) and the readers' performances on all three comprehension measures indicates the association between using this strategy (Pause) and less-successful reading outcomes.

Consistent with Endley's (2016) observation, the participants in this study were noted re-reading or translating what they had just read during their pauses. For example, when asked what they were thinking when they paused in the interview, the participant's response was:

> "I was thinking of going back to see the first and last sentence of the paragraph because I wanted to summarize it. But I realized that I could not summarize it because it was too difficult. Then, I continued to read on." (Participant A)

> "Oh, this is when I realized that I did not know many of the words. Then, I reached for my phone (used as a dictionary)." (Participant E)

> "Probably because the sentence was too long for me to understand, I was thinking about re-reading it." (Participant F)

While appearing to be doing nothing, when pausing, on the camera, the participants were clearly monitoring their comprehension breakdown and actively employing other strategies such as re-reading or translation silently to solve these problems. However, their realization of comprehension breakdown and intention to bridge the comprehension gap did not help these participants achieve better results. One explanation is that the readability of the texts was low. Consequently, the participants did not have the linguistic knowledge to solve their realized comprehension problems. As a result, their effort was futile.

Readers paused because they did not comprehend what they had just read. Grabe and Stoller (2011) argued that, for fluent readers, lower-level linguistic processes (e.g., word recognition, sentence parsing) are carried out relatively automatically without being noticed. However, when some aspects of these processes break down, the reader then becomes aware of the problems and pauses to extract the most appropriate meaning from the reading. The negative association between Prob5 and the readers' performances corroborates Grabe and Stoller's argument and highlights the inherent role of the linguistic knowledge base for both the flow of lower-level decoding processes and the ultimate comprehension in L2 reading.

Let us now turn to the nine strategies that contributed greatly to the readers' performance on at least one of the measures. To reiterate, successful readers in this study used more of the following nine strategies: Sup1 (TakeNotes), Sup3 (UseDictionary), Sup7 (AskSelfQs), Sup8 (Translate), Glob5 (UseTables), Glob7 (TypographicalFeatures), Glob10 (Predict), Glob11 (CheckPrediction), and Prob10 (AnalyzeGrammar). Note that six of these nine strategies (Glob5, Glob7, Glob10, Glob11, Sup1, and Sup7) work to facilitate the reader's holistic comprehension of the texts. This confirms that readers with better reading outcomes tend to emphasize the overall textual meaning. This finding is consistent with prior literature in the general agreement that successful readers tend to use more top-down strategies (e.g., questioning content) that facilitate global comprehension of the text, whereas less-successful readers favor bottom-up strategies (e.g., translation, underling) that target word and sentence-level decoding (Block, 1986; Brantmeier, 2002; Carrell, 1989; Lin & Yu, 2013; Wang, 2016).

The claim "less successful readers favored bottom-up strategies" does not necessarily indicate that these strategies caused less-successful outcomes. As evidenced in our data, three bottom-up strategies (Sup3 "UseDictionary", Sup8 "Translation", Prob10 "AnalyzeGrammar") had a great contribution to the readers' better comprehension outcomes. This finding is not at all surprising given the fact that the text was relatively difficult to read. The challenges related to vocabulary and grammar required readers to carry out appropriate decoding strategies to achieve better lower-level linguistic processing outcomes. These served as building blocks for higher-level comprehension processes, such as constructing a text model of comprehension (Grabe and Stoller, 2011). We caution researchers and practitioners against the temptation to link less successful readers with decoding strategies because the effectiveness of a strategy is context dependent and should not be generally labeled (Cohen, 1986; Yang, 2006).

RQ3: With verbal reports, what is the relationship between the frequency and variety of strategy use and reading comprehension, as assessed with free recall, sentence completion, and multiple-choice?

We found that frequency was not associated with participants' performance on any of the three comprehension measures, and that variety only contributed to participants' FR score. This finding contradicts Anderson (1991) and Wang's (2016) result that participants employed similar types of reading strategies (variety) regardless of their performance on multiple-choice questions. The conflicting results may be due to the difference in the L2 reading comprehension measures involved (FR in the current study as opposed to MC in Anderson and Wang's investigation). This again suggests that comprehension test type matters for interpreting the relationship between strategy use and reading comprehension. Hence, to capture a more complete picture of how strategy use associates with different reading outcomes, a combination of various comprehension measures should be used in L2 reading strategy research.

Implications for L2 Reading Strategy Instruction

These findings provide a few implications for researchers and practitioners. First, the "gap" between the number of strategies that readers were aware of and that they actually employed suggests that being strategic is not simply a matter of knowing which strategies to use, but also when, why, and how to use them effectively (Anderson, 1991; Cohen, 2014). To help L2 readers bridge the gap between "knowing what" and "knowing how", and subsequently improve their comprehension, explicit strategy instructions are recommended. These types of instructions have been found to be worthy of the time and effort (Cohen, 2014; Macaro & Erler, 2008; Schueller, 2004; Yapp et al., 2021a; Yapp et al., 2021b). For example, in a recent meta-analysis of 46 studies on the effectiveness of strategy training on L2 reading comprehension outcomes, Yapp, De Graaff and Van den Bergh (2021b) found a large effect size (g = 0.91), confirming the effectiveness of strategy training on L2 reading comprehension. Additionally, our data showed that a greater variety of strategy use promoted better reading performance, which adds support to Pressley (2002) and Guthrie and Humenick's (2004) argument that reading strategy instructions should consider the full variety of strategies available.

Second, Cohen (2014) pointed out that "there is no one best way to conduct strategy instruction–it depends on the particular learning context and the given learners." (p.120) Our data suggest that the reader's socio-cultural background and instructional experience need to be taken into consideration when developing strategy training programs or instructions. For example, Chinese EFL learners' reading habits are persistently reinforced by the washback effect. To counteract this unrelenting influence and help Chinese EFL learners become strategic readers, explicit strategy instruction in a "small but consistent" fashion is crucial (Grabe & Stoller, 2011; Macaro & Erler, 2008). For instance, teachers could incorporate a 20-minute explicit strategy instruction into every lesson during the semester. This way, effective strategy use could be constantly strengthened.

Third, while teaching L2 readers to be more strategic, practitioners should devote time and instruction to increase L2 readers' linguistic knowledge, including both vocabulary and

grammar. This is particularly crucial for the readers' comprehension of domain-specific texts in a target language, as the terminology in a given field is often used less frequently and is thus less likely to be acquired incidentally. Participants in the current study clearly stated that they had problems recognizing the concepts they had known due to a lack of vocabulary. For example, they knew the concept of "hypertension" but did not know the word in English. As a result, readers devoted a great deal of time and cognitive resources to figuring out the meaning of the unfamiliar words by employing decoding strategies such as using a dictionary and grammatical analyses. Subsequently, they neglected the overall textual meaning and their comprehension suffered. Vocabulary and grammar knowledge are the linguistic foundation on which readers build their reading strategy skills, which foster automatic decoding processes and better reading comprehension outcomes (Anderson, 1991; Grabe & Stoller, 2011). Explicit strategy instructions coupled with growth in linguistic knowledge may help build strategic readers and possibly lead to success in their fields.

Last, our data clearly indicate that the contribution of reading strategies to reading comprehension varied when different comprehension tasks are used. This finding suggests that, when evaluating the effectiveness of strategy instruction on reading performance, a variety of comprehension tasks is necessary for a more accurate assessment.

Conclusion

Our data reveal that university-level Chinese EFL readers of higher and lower proficiency levels favored six decoding strategies while reading two domain-specific texts. The discrepancy between the survey and verbal reports for assessing strategy use suggests that readers in the study are aware of a wide range of reading strategies; however, in practice participants either ignore the majority of the strategies or do not use them to the extent to which they "think" they do. There is a clear gap between "what they know" and "what they do" in terms of strategy use in L2 reading. We recommend that practitioners bridge the gap through explicit strategy instruction. In addition, while our data support the idea that readers of higher reading ability employ more strategies facilitating the holistic comprehension of texts, the results also highlight the fundamental role of decoding strategies. Finally, readers who have better comprehension outcomes tend to employ a greater variety of strategies. These findings suggest that practitioners should include a wide range of reading strategies in their strategy instruction.

Moreover, our findings indicate that comprehension task type moderates the relationship between strategy usage and reading comprehension and needs to be examined as an independent variable in future research. To obtain a more thorough understanding of how strategy use associates with reading performance, a combination of different assessment tasks should be considered in a single study. We provide practitioners with important implications on L2 reading strategy training, such as considering the readers' socio-cultural background and instructional experiences when designing strategy instruction, promoting readers' linguistic knowledge growth, and utilizing various reading comprehension tasks when evaluating the effectiveness of a strategy training program.

It is noteworthy that due to the qualitative nature of verbal reports, the sample size in this study is relatively small (ten participants). Thus, the empirical findings are largely limited to this

sample group. Nevertheless, our results call attention to several variables (e.g., differences between reading strategy measures, comprehension task type) that are crucial in L2 reading strategy research and which warrant further investigations with larger sample sizes.

References

Akyel, A., & Erçetin, G. (2009). Hypermedia reading strategies employed by advanced learners of English. *System*, *37*, 136-152.

Alderson, J. C. (2000). *Assessing reading*. Cambridge University Press.

Alkhaleefah, Tarek A. (2017). Saudi EFL learners' reported reading problems and strategic processing of text types: A think-aloud case study. *Reading Psychology*, *38*, 687-730.

Alsheikh, N. O., & Mokhtari, K. (2011). An examination of the metacognitive reading strategies used by native speakers of Arabic when reading in English and Arabic. *English Language Teaching*, *4*, 151-160.

Anderson, N. J. (1991). Individual differences in strategy use in second language reading and testing. *The Modern Language Journal*, *75*, 460-472.

Bakhshalinezhad, L., Nikou, F. R., & Bonyadi, A. (2015). Using the think-aloud technique for determining different reading strategies used by Iranian EFL learners. *Advances in Language and Literary Studies*, *6*, 15-22.

Bang, H., & Zhao, C. (2007). Reading strategies used by advanced Korean and Chinese ESL graduate students: A case study. *The Reading Matrix*, *7*, 30-50.

Bernhardt, E.B. (2011). *Understanding advanced second-language reading*. Routledge.

Block, E. L. (1986). The comprehension strategies of second language readers. *TESOL Quarterly*, *20*, 463–494.

Brantmeier, C. (2002). Second language reading strategy research at the secondary and university levels: Variations, disparities, and generalizability. *The Reading Matrix*, *2*, 1-14.

Brantmeier, C. (2005). Effects of reader's knowledge, text type, and test type on L1 and L2 reading comprehension in Spanish. *The Modern Language Journal*, *89*, 37-53.

Brantmeier, C. (2006). Advanced L2 learners and reading placement: Self-assessment, CBT, and subsequent performance. *System*, *34*, 15-35.

Brantmeier, C., & Dragiyski, B. (2009). Toward a dependable measure of metacognitive reading strategies with advanced L2 learners. In Brantmeier, C. (Ed.), *Crossing languages and research methods: Analyses of adult foreign language reading* (pp. 47-72). IAP.

Brantmeier, C., Strube, M. & Yu, X. (2014). Scoring recalls for L2 readers of English in China: Pausal or idea units. *Reading in a Foreign Language*. *26*, 114-130.

Carrell, P. L. (1989). Metacognitive awareness and second language reading. *The Modern Language Journal*, *73*, 121-134.

Chern, C. (1994). Chinese readers' metacognitive awareness in reading Chinese and English. In Bird, N. (Ed., and others). Papers presented at the Annual International Language in Education Conference. Hong Kong.

Cohen, A. D. (2014). *Strategies in learning and using a second language*. Routledge.

Durning, S. J., Artino Jr., A. R., Beckman, T. J., Graner, J., van der Vleuten, C., Holmboe, E., & Schuwirth, L. (2013). Does the think-aloud protocol reflect thinking? Exploring functional neuroimaging differences with thinking (answering multiple choice questions) versus thinking aloud. *Medical Teacher*, *35*, 720-726.

Endley, Martin J. (2016). Proficiency as a variable in Gulf EFL students' employment of reading strategies. *Reading in a Foreign Language*, *28*, 183-223.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data*. Massachusetts Institute of Technology.

Ferguson, C. J. (2009). An effect size primer: A guide for clinicians and researchers. *Professional Psychology: Research and Practice, 40*, 532-538.

Gass, S. M., & Mackey, A. (2013). *Stimulated recall methodology in second language research*. Routledge.

Ghavamnia, M., Ketabi, S., & Tavakoli, M. (2013). L2 reading strategies used by Iranian EFL learners: A think-aloud study. *Reading Psychology*, *34*, 355-378.

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

Grabe, W., & Stoller, F. L. (2011). *Teaching and researching reading*. Routledge.

Guthrie, J. T. and Humenick, N. (2004) Motivating students to read: Evidence for classroom practices that increase reading motivation and achievement. In P. McCardle and V. Chhabra (Eds.), *The voice of evidence in reading research* (pp. 329– 54). Paul H. Brookes.

Han, F. (2017). Applicability of the Compensatory Encoding Model in foreign language reading: An investigation with Chinese college English language learners. *Frontiers in Psychology*, *8*, 681-693.

Han, F., & Stevenson, M. A. R. I. E. (2008). Comprehension monitoring in first and foreign language reading. *University of Sydney Papers in TESOL*, *3*, 73-110.

He, T. (2008). Reading for different goals: The interplay of EFL college students' multiple goals, reading strategy use and reading comprehension. *Journal of Research in Reading*, *31*, 224-242.

Hosenfeld, C. (1977). A preliminary investigation of the reading strategies of successful and nonsuccessful second language learners. *System*, *5*, 110–123.

Koda, K. (2005). *Insights into second language reading: A cross-linguistic approach*. Cambridge University Press.

Kong, A. (2006). Connections between L1 and L2 readings: Reading strategies used by four Chinese adult readers. *The Reading Matrix*, *6*, 19-45.

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

Lin, L., & Yu, W. (2015). A think-aloud study of strategy use by EFL college readers reading Chinese and English texts. *Journal of Research in Reading, 38*, 286-306.

Lipsey, M. W., Puzio, K., Yun, C., Hebert, M. A., Steinka-Fry, K., Cole, M. W., Roberts, M., Anthony, K. S., & Busick, M. D. (2012). Translating the statistical representation of the effects of education interventions into more readily interpretable forms. *National Center for Special Education Research*.

Liu, H., & Brantmeier, C. (2019). "I know English": Self-assessment of foreign language reading and writing abilities among young Chinese learners of English. *System*, *80*, 60-72.

Macaro, E., & Erler, L. (2008). Raising the achievement of young-beginner readers of French through strategy instruction. *Applied Linguistics*, *29*, 90-119.

Malcolm, D. (2009). Reading strategy awareness of Arabic-speaking medical students studying in English. *System*, *37*, 640-651.

Mokhtari, K., & Reichard, C. A. (2002). Assessing students' metacognitive awareness of reading strategies. *Journal of Educational Psychology*, *94*, 249-259.

Mokhtari, K., & Reichard, C. (2004). Investigating the strategic reading processes of first and second language readers in two different cultural contexts. *System*, *32*, 379-394.

Mokhtari, K., & Sheorey, R. (2002). Measuring ESL students' awareness of reading strategies. *Journal of Developmental Education*, *25*, 2-11.

Pressley, M. (2002) Metacognition and self-regulated comprehension. In A. Farstrup and S. Samuels (Eds.), *What research has to say about reading instruction* (pp. 291–309). International Reading Association.

Pritchard, R. (1990). The effects of cultural schemata on reading processing strategies. *Reading Research Quarterly*, *25*, 273-295.

Schueller, J. (2004). Gender and foreign language reading comprehension: The effects of strategy training. In Brantmeier, C. (Ed.), *Crossing languages and research methods: Analyses of adult foreign language reading* (pp. 147-167). IAP.

Sheorey, R., & Mokhtari, K. (2001). Differences in the metacognitive awareness of reading strategies among native and non-native readers. *System*, *29*, 431-449.

Taki, S. (2016). Metacognitive online reading strategy use: Readers' perceptions in L1 and L2. *Journal of Research in Reading*, *39*, 409-427.

Tsai, Y., Ernst, C., & Talley, P. (2010). L1 and L2 strategy use in reading comprehension of Chinese EFL readers. *Reading Psychology*, *31*, 1-29.

Wang, Y. (2016). Reading strategy use and comprehension performance of more successful and less successful readers: A think-aloud study. *Educational Sciences: Theory and Practice*, *16*, 1789-1813.

Yang, Y. F. (2006). Reading strategies or comprehension monitoring strategies? *Reading Psychology*, *27*, 313-343.

Yapp, D., de Graaff, R., & van den Bergh, H. (2021). Effects of reading strategy instruction in English as a second language on students' academic reading comprehension. *Language Teaching Research*. 74-80. https://doi.org/10.1177/1362168820985236

Yapp, D., de Graaff, R., van den Bergh, H. (2021). Improving second language comprehension through reading strategies: A meta-analysis of L2 reading strategy interventions. *Journal of Second Language Studies*, *4*, 154-192.

Yayli, D. (2010). A think-aloud study: Cognitive and metacognitive reading strategies of ELT department students. *Eurasian Journal of Educational Research*, *38*, 234-251.

Zhang, L., & Wu, A. (2009). Chinese senior high school EFL students' metacognitive awareness and reading-strategy use. *Reading in a Foreign Language*, *21*, 37-59.

Zhou, X., & Zhao, Y. (2014). A comparative study of reading strategies used by Chinese English majors. *English Language Teaching*, *7*, 13-18.

Appendix A

Demographic Questionnaire

Please complete the following:
1.   Gender:
2.   Age:
3.   Major area of study:
4.   Native language(s):
5.   Are you a _____ (first year, second year, third year, fourth year, graduate) student?
6.   How many hours of English classes do you attend every week? _____
7.   Where do you plan to use English in the future?
_____
8.   Are you planning to go to graduate school to study medicine? (please circle one)
        Yes   Maybe   No
9.   How many years have you studied English? _____
10.   How many hours on average do you spend reading in English every week? _____
11.   What percentage of readings about medicine do you read in English each week?
        _____
12.   Do you think it is important to learn English well? (please circle one from below)
        Absolutely yes   not sure   no
        Why?   Or   why   not?
        _____
        _____
        ____
13.   How interested are you in learning English? (please circle one)
        very interested   a little interested   neutral   not very interested   not interested at all
14.   Ever since you started to learn English, how often have you been taught to use reading strategies? (please circle one)
        Never   seldom   sometimes   often   very often
15.   How familiar are you with different types of reading strategies? (please circle one)
        Not familiar at all   somewhat unfamiliar   neutral   familiar   very familiar
16.   Which CET have you taken_____, and what is your overall score on that? _____, and what is the score for the reading section? _____
17.   How long ago did you take the CET (please circle one):
        within 6 months, within 1 year, within 1.5years, within 2 years, more than two years ago

Appendix B

Sample Reading Passages, Free Recall, Sentence Completion and Multiple-Choice Questions

Test 1:
Please read the passage following the TALK ALOUD instruction. You CAN use a dictionary while reading. After reading the passage, you will also complete three reading comprehension tasks WITHOUT LOOKING BACK AT THE PASSAGE. Please note that this passage is taken from an article published in the journal of *Perspective in Biology and Medicine*.

Modern clinical applications related to Chinese traditional theories of drug interactions

Conclusions

Even though the Chinese theories on drug action have become dated, the time frame has not been that long. Pharmacology did not emerge as a disciplinary entity in the West until the end of the 19th century, and only then did the textbooks begin to give attention to the factors that may alter drug action, which the Chinese had thought about and written about centuries earlier. What the Chinese enunciated with respect to drug action transcends the thousands of remedies they discovered for treating the ills of their population. What they conjured and anticipated parallels in many respects the modern concepts of the West. Their theories, however, lack the flexibility to incorporate the rapid advances made in modern science. In contrast to the Western approach, wherein new findings can be applied to modify or correct existing theories, with traditional Chinese medicine, new data must be incorporated into existing dogma that is immutable. As a consequence, a curious dichotomy exists today in traditional medicine between its basic and practical facets. While pharmaceutical chemists and pharmacologists strive to isolate and characterize active plant principles, herbalists throughout the world, including Europe and the United States in addition to Asia, continue to ply their trade applying outmoded ancient theories and methodologies.

Despite the criticisms of the traditional theories, the beneficial consequences for mankind have been many and varied. It is irrefutable that a large number of useful preparations were discovered at the time the notions were in effect. No doubt the discovery of useful natural products in the beginning was empiric and serendipitous, but the concepts formulated later from such findings were essentially a concept unique to Chinese civilization. The fruits from the *yin-yang* and five elements doctrine include efficacious medicinal remedies still in use. Undoubtedly, the greatest gift the Chinese contributed to preserving the health of the human race is the first immunologic procedure, smallpox inoculation. The eminent historian Joseph Needham points out this was the beginning of immunology, the most beneficent department of modern medical science. Besides the therapeutic agents, the holistic prerequisites brought into focus the importance of preventive medicine, calisthenics, and massage that are so popular in modern health facilities. The theories facilitated the systematization and rationalization of a large mass of data on medicinal remedies, the development of proto-sciences that have become important basic and clinical disciplines related to pharmaceutics, and the compilation of invaluable compendia on *materia medica*, including the first pharmacopeia and the formulation of some principles in pharmacology that can be extrapolated for current application.

Free recall. Without looking back at the passage, recall <u>in Chinese</u> as much as you can of what you just read. Try to recall main ideas as well as details. The emphasis is on the quantity recalled.

Sentence completion. Based on the text you just read, please complete the following sentences. You can use Chinese to complete the sentences if you want.
Sample 1: Use one word or phrase to describe the author's overall attitude towards Chinese traditional theories on medicine: _____.
Sample 2: Traditional Chinese medicine _____
incorporating the rapid advances made in modern science.

Multiple-choice questions. Based on the text you just read, please circle the letter of the best answer to each of the following questions (there is only ONE correct answer).
Sample 1: Which of the following words or phrases best describes the author's attitude towards Chinese theories on drug action?
a.   Criticizing
b.   Overall positive
c.   Neutral
d.   Overall negative
Sample 2: Which of the following statements is NOT true?
e.   Chinese had thought about and written about the factors that may alter drug action centuries earlier before Pharmacology emerged as a disciplinary entity in the West.
f.   Centuries earlier before Pharmacology emerged as a disciplinary entity in the West, the Chinese had already discovered thousands of remedies for treating the ills of their population.
g.   Traditional Chinese medicine is flexible with incorporating the rapid advances made in modern science.
h.   There is a dichotomy existing today in traditional medicine between its basic and practical facets.
Text 2
Please read the passage following the TALK ALOUD instruction. You CAN use a dictionary while reading. After reading the passage, you will also complete three reading comprehension tasks WITHOUT LOOKING BACK AT THE PASSAGE. Please note that this passage is taken from an article published in the journal of *Perspective in Biology and Medicine*.

Control of Cardiovascular Disease in the 20th Century: Meeting the Challenge of Chronic Degenerative Disease.
Making It Happen: Cardiovascular Disease (CVD) as a Public Health Priority
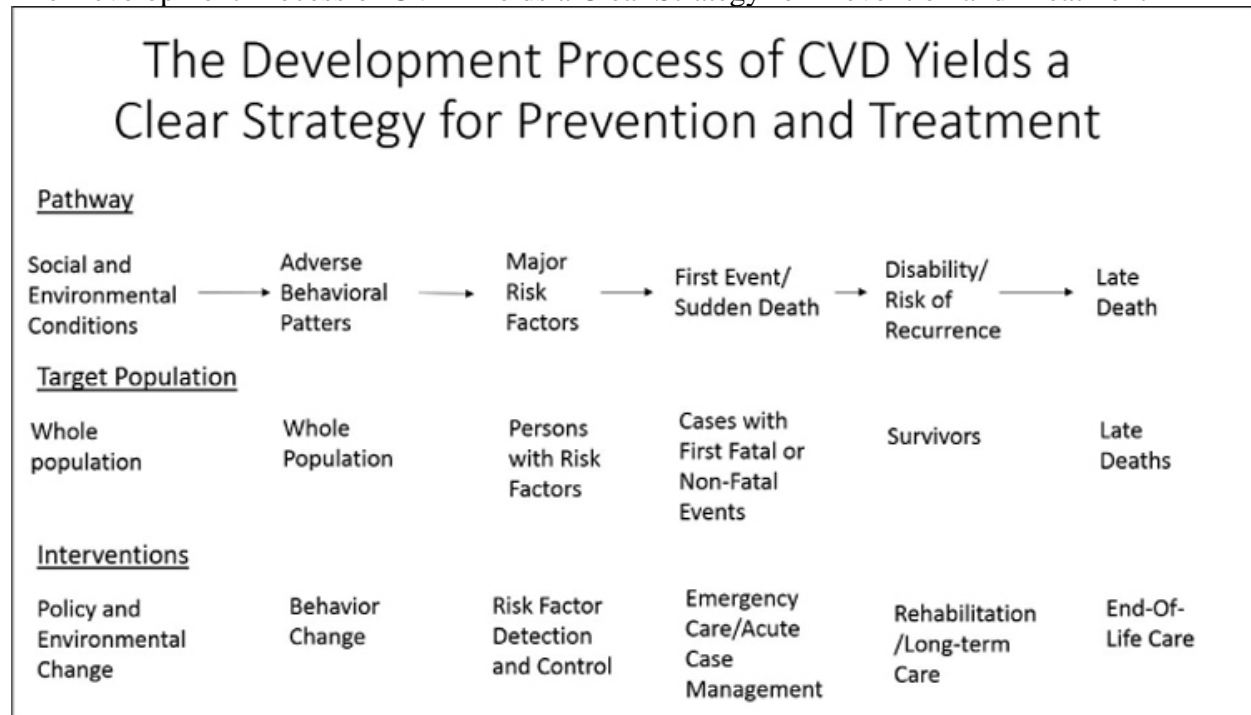A coalition of professional organizations, most prominently the American Heart Association, and federal agencies, with the National Heart, Lung and Blood Institute (NHLBI) in the lead, launched a vigorous public education campaign aimed at the general public, patients, and physicians. The NHLBI played a key role through the National High Blood Pressure Education Program (NCEP), which promoted screening, detection, and control with pharmacologic agents, and the National Cholesterol Education Program, which initially created awareness of the importance of high cholesterol, disseminated information on beneficial dietary patterns, and subsequently set guidelines for drug treatment.
National survey data from the National Health and Nutrition Examination Survey (NHANES) documented the decline in mean serum cholesterol from 220 mg/dl in the 1980s to less than 200 mg/dl at the present time, and the increase in treatment and control of hypertension to greater than 50% in the population. The advent of statins in the 1990s led to a major improvement in cholesterol control in high-risk individuals. Over this entire period, of course, advances in

hospital methods led to higher survival rates with myocardial infarction and continuous improvement in surgical and catheter-based methods to re-vascularize diseased coronary and carotid vessels. A well-developed strategy now exists to intervene on every stage of the etiologic sequence from dietary intake to rehabilitation and secondary prevention (Table 1).

Of course, all of these efforts were met with staunch resistance. The food industry, particularly the egg and beef interests, fought tooth and nail to delay public health action. Scientific evidence was criticized, the reputations of individual scientists were impugned, data were subpoenaed for "re-analysis," and a small cohort of academics allowed themselves to be used as "responsible critics." These efforts certainly sowed confusion in the media and created a legacy of hucksters and quacks who to this day promote a panoply of theories and diets that lack scientific support. These disruptive voices have no doubt limited progress, at the cost of many needless deaths and lives of long-term disability. But this level of dissension, if so it can be called, is inevitable in a pluralistic society. In truth it must be said that the CVD prevention message has demonstrated remarkable durability and success. The positive outcomes must be taken as a tribute to the elegance and validity of the theory and the dedication and skill of those who have fought to make a heart healthy lifestyle the norm, and as proof—most of all—that in the long-term success breeds ever greater success.

Table 1.
The Development Process of CVD Yields a Clear Strategy for Prevention and Treatment

The Development Process of CVD Yields a Clear Strategy for Prevention and Treatment

Pathway

| Social and Environmental Conditions | → | Adverse Behavioral Patters | → | Major Risk Factors | → | First Event/ Sudden Death | → | Disability/ Risk of Recurrence | → | Late Death |

Target Population

| Whole population | | Whole Population | | Persons with Risk Factors | | Cases with First Fatal or Non-Fatal Events | | Survivors | | Late Deaths |

Interventions

| Policy and Environmental Change | | Behavior Change | | Risk Factor Detection and Control | | Emergency Care/Acute Case Management | | Rehabilitation /Long-term Care | | End-Of-Life Care |

Free recall. Without looking back at the passage, recall <u>in Chinese</u> as much as you can of what you just read. Try to recall main ideas as well as details. The emphasis is on the quantity recalled.

Sentence completion. Based on the text you just read, please complete the following sentences. You can use Chinese to complete the sentences if you want.

Sample 1: The American Heart Association, and federal agencies, with the National Heart, Lung and Blood Institute (NHLBI) in the lead,

_____ aimed at the general public, patients, and physicians.

Sample 2: The National survey data from the National Health and Nutrition Examination Survey (NHANES) documented that (please name at least one)

_____ .

Multiple-choice questions. Based on the text you just read, please circle the letter of the best answer to each of the following questions (there is only ONE correct answer).

Sample 1: What did the American Heart Association and the National Heart, Lung and Blood Institute (NHLBI) do to help fight cardiovascular disease (CVD) and make it a public health priority?

a.   Trained physicians
b.   Told the public about the best food for fighting CVD
c.   Launched a vigorous public education campaign
d.   Provided the public with effective treatment

Sample 2: Which of the following phenomena was documented by the National survey data from the National Health and Nutrition Examination Survey (NHANES)?

a.   The success in treating diabetes
b.   The increase in mean serum cholesterol

c.   The success in treating high blood pressure
d.   The success in treating heart attack

Appendix C

Survey of Reading Strategies for L2 Contexts

The purpose of the survey is to collect information about the various techniques you use when you read academic materials in English (e.g., reading textbooks for homework or examinations, reading journal articles, etc.).

All the items below refer to your reading of college-related academic materials (such as textbooks, *not* newspapers or magazines). Each statement below is followed by five numbers, 1, 2, 3, 4, and 5, and each number means the following:

"1" means that 'I never or almost never do this'.
"2" means that 'I do this only occasionally'.
"3" means that 'I sometimes do this'. (About 50% of the time.)
"4" means that 'I usually do this'.
"5" means that 'I always or almost always do this'.

After reading each statement, *circle the number* (1,2,3,4, or 5) which applies to you. Note that there are no right or wrong responses to any of the items on this survey.

	TYPE	STRATEGIES	SCALE				
1	SUP	I take notes while reading to help me understand what I read.	1	2	3	4	5
2	GLOB	I think about what I know to help me understand what I read.	1	2	3	4	5
3	GLOB	I take an overview of the text to see what it is about before reading it.	1	2	3	4	5
4	PROB	I read slowly and carefully to make sure I understand what I am reading.	1	2	3	4	5
5	GLOB	I skim the text first by noting its characteristics like length and organization.	1	2	3	4	5
6	PROB	I try to get back on track when I lose concentration.	1	2	3	4	5
7	SUP	I underline or circle information in the text to help me remember it.	1	2	3	4	5
8	PROB	I adjust my reading speed according to what I am reading.	1	2	3	4	5
9	GLOB	When reading, I decide what to read closely and what to ignore.	1	2	3	4	5
10	SUP	I use reference materials (e.g., a dictionary) to help me understand what I read.	1	2	3	4	5
11	PROB	When text becomes difficult, I pay closer attention to what I am reading.	1	2	3	4	5
12	GLOB	I use tables, figures, and pictures in the text to increase my understanding.	1	2	3	4	5
13	PROB	I stop from time to time and think about what I am reading.	1	2	3	4	5

14	GLOB	I use context clues to help me better understand what I read.	1	2	3	4	5
15	SUP	I paraphrase (restate ideas in my own words) to better understand what I read.	1	2	3	4	5
16	PROB	I try to picture or visualize information to help remember what I read.	1	2	3	4	5
17	SUP	I summarize what I read to reflect on important information in the text.	1	2	3	4	5
18	GLOB	I use typographical features like boldface and italics to identify key information.	1	2	3	4	5
19	GLOB	I critically analyze and evaluate the information presented in the text.	1	2	3	4	5
20	SUP	I go back and forth in the text to find relationships among ideas in it.	1	2	3	4	5
21	GLOB	I check my understanding when I come across new information.	1	2	3	4	5
22	GLOB	I try to guess what the content of the text is about when I read (making predictions).	1	2	3	4	5
23	PROB	When text becomes difficult, I re-read it to increase my understanding.	1	2	3	4	5
24	SUP	I ask myself questions I like to have answered in the text.	1	2	3	4	5
25	GLOB	I check to see if my guesses about the text are right or wrong.	1	2	3	4	5
26	PROB	When I read, I guess the meaning of unknown words or phrases.	1	2	3	4	5
27	SUP	When reading, I translate from English into my first language.	1	2	3	4	5
28	SUP	When reading, I think about information in both English and my mother tongue.	1	2	3	4	5
29	GLOB	When reading, I tend to identify main ideas and distinguish them from supporting ideas.	1	2	3	4	5
30	GLOB	I tend to identify the text type and the structure or discourse patterns of the text.	1	2	3	4	5
31	PROB	When reading, I purposely ignore some unknown words and read on.	1	2	3	4	5
32	GLOB	While reading, I constantly check if I know the main ideas of the text and clearly know it when there is a breakdown.	1	2	3	4	5
33	PROB	I analyze grammatical structures (e.g., conjunctive adverbs and clauses) to help me understand the text.	1	2	3	4	5

Appendix D

The Coding Scheme for Yes/No Strategies (19 out of 33) and Countable Strategies (14 out of 33)

Type	Identifier	Reading strategies
YesNo strategies	Back&Forth	Sup6: I go back and forth in the text to find relationships among ideas in it.
	ThinkL1&L2	Sup9: When reading, I think about information in both English and my mother tongue.
	PriorKnowledge	Glob1: I think about what I know to help me understand what I read.
	Overview	Glob2: I take an overview of the text to see what it is about before reading it.
	SkimFirst	Glob3: I skim the text first by noting its characteristics like length and organization.
	What2Ignore	Glob4: When reading, I decide what to read closely and what to ignore.
	UseTables	Glob5: I use tables, figures, and pictures in the text to increase my understanding.
	TypographicalFeatures	Glob7: I use typographical features like boldface and italics to identify key information.
	NewInfo	Glob9: I check my understanding when I come across new information.
	DiscoursePtns	Glob13: I tend to identify the text type and the structure or discourse patterns of the text.
	FollowIdeas	Glob14: While reading, I constantly check if I know the main ideas of the text and clearly know it when there is a breakdown.
	CarefulReading	Prob1: I read slowly and carefully to make sure I understand what I am reading.
	AdjustSpeed	Prob3: I adjust my reading speed according to what I am reading.
	SlowWhenDiff	Prob4: When text becomes difficult, I pay closer attention to what I am reading.
	VisualizeInfo	Prob6: I try to picture or visualize information to help remember what I read.
	Re-read	Prob7: When text becomes difficult, I re-read it to increase my understanding.
	GuessMeaning	Prob8: When I read, I guess the meaning of unknown words or phrases.
	IgnoreUnknown	Prob9: When reading, I purposely ignore some unknown words and read on.
	AnalyzeGrammar	Prob10: I analyze grammatical structures (e.g.,

		conjunctive adverbs and clauses) to help me understand the text.
Countable strategies	TakeNotes	Sup1: I take notes while reading to help me understand what I read.
	Underline	Sup2: I underline or circle information in the text to help me remember it.
	UseDictionary	Sup3: I use reference materials (e.g., a dictionary) to help me understand what I read.
	Paraphrase	Sup4: I paraphrase (restate ideas in my own words) to better understand what I read.
	Summarize	Sup5: I summarize what I read to reflect on important information in the text.
	AskSelfQs	Sup7: I ask myself questions I like to have answered in the text.
	Translate	Sup8: When reading, I translate from English into my first language.
	ContextClues	Glob6: I use context clues to help me better understand what I read.
	EvalInfo	Glob8: I critically analyze and evaluate the information presented in the text.
	Predict	Glob10: I try to guess what the content of the text is about when I read (making predictions).
	CheckPrediction	Glob11: I check to see if my guesses about the text are right or wrong.
	MainVsSupport	Glob12: When reading, I tend to identify main ideas and distinguish them from supporting ideas.
	LoseConcent	Prob2: I try to get back on track when I lose concentration.
	Pause	Prob5: I stop from time to time and think about what I am reading.

About the Authors

Yanjie Li, the corresponding author, is a PhD candidate studying applied linguistics at Washington University in St. Louis. Her research interests involve language learning and teaching in EFL and ESL contexts. Yanjie is currently working on exploring factors that affect L2 learner's reading comprehension such as reading strategies, working memory, and adjunct questions.
E-mail: li.yanjie@wustl.edu

Cindy Brantmeier is a professor of Applied Linguistics in Global Studies and Faculty Fellow for International Research in the Office of the Vice Chancellor for Research at Washington University in St. Louis. Her research examines variables involved in second language reading, language testing and assessment and health literacy.
E-mail: cbrantme@wustl.edu

Yanming Gao is an associate professor of Applied Linguistics at the Training Center of the Ministry of Education for Studying Overseas at Northeast Normal University. Her published work focuses on English reading, speaking, language teaching and learning methodology.
E-mail: gaoym192@nenu.edu.cn

Mark Hogrebe is an educational researcher in the Department of Education at Washington University in St. Louis. His interests include research and evaluation methodologies in applied settings, education in the STEM fields, diversity issues, and using geographic information systems to give geospatial perspective to education data.
E-mail: mhogrebe@wustl.edu