
Modeling cognitive and metacognitive strategies and their relationships to EFL reading test performance

Aek Phakiti

University of Sydney

Abstract

This article reports on an empirical study that examines the nature of cognitive strategies (comprehending, retrieval and memory strategies) and metacognitive strategies (planning, monitoring and evaluating strategies) and their direct and indirect relationships to English as a foreign language (EFL) reading test performance, employing the structural equation modeling (SEM) approach. The study was carried out at a government university in Thailand in which 358 students took a reading comprehension test and immediately after completing it, answered a questionnaire on their strategy use. The SEM results show that: (1) memory and retrieval strategies facilitated EFL reading test performance via comprehending strategies; (2) monitoring strategies performed an executive function on memory strategies, whereas evaluating strategies regulated retrieval strategies; (3) planning strategies did not directly regulate memory, retrieval or comprehending strategies, but instead regulated these cognitive strategies via monitoring and evaluating strategies; and (4) only comprehending strategies were found to directly influence EFL reading test performance.

Address for correspondence: Aek Phakiti, TESOL Program, Faculty of Education and Social Work, University of Sydney, NSW 2006, Australia; Email: a.phakiti@edfac.usyd.edu.au

Melbourne Papers in Language Testing, 1, 53-95.
ISSN 1327-0311 (print)

1. Introduction

The present article investigates cognitive and metacognitive strategy use in response to reading test tasks. There are two major objectives of this article. The first is to present an empirical model of the complex relationships between cognitive and metacognitive strategies and L2 reading performance. The second is to demonstrate how the structural equation modeling (SEM) methodology can provide insights into such complex relationships. This particular methodological approach is important but has not been employed much in L2 reading and testing research. The SEM approach can analyze the structure and effect of unobservable, latent traits through the analysis of interindividual differences data, by statistically relating covariation between observed variables to latent variables.

2. Review of literature

L2 reading

Reading in a second language (L2) is complex, dynamic and multi-dimensional (Alderson, 2000). This is primarily because it involves interactions among the reader's interlanguage competence (e.g., incomplete, fragmented or not fully-developed linguistic, strategic, discourse and sociolinguistic competence), personal characteristics (e.g., learning and cognitive style, gender, motivation and volition, socioeconomic status, educational levels) and external contexts (e.g., topics, text characteristics, reasons to read, stakes of reading, time constraints).

The past decades have seen research into the nature of L2 reading in which an attempt has been made to understand its nature by investigating reader factors (e.g., L1 literacy, language proficiency, background knowledge, knowledge of genre and pragmatics, metalinguistic knowledge, metacognition, strategy use, motivation and affect) and contextual factors (e.g., text topic and content, text type and genre, text readability, verbal and non-verbal communication; see e.g., Alderson, 2000). Of these factors, the present article focuses specifically on the nature of cognitive and metacognitive strategies and their relationships to English as a foreign language (EFL) reading comprehension.

Levels of reading processing

When individuals are reading, their reading processes would range from lower-level to higher-level processing (Alderson, 2000; Kintsch, 1998; Pressley & Afflerbach, 1995). Lower-level processing includes automatic recognition of word meanings, syntactic structures and parts of speech. Automatic reading processing suggests that individuals may read with a minimum of awareness. For example, readers can decode print without really thinking about it. Hence, part of success in L2 reading depends on the level of automaticity (Segalowitz, 2003). Given the complexity of text, the more automatized the readers' processing is, the more efficient reading will be (Alderson, 2000). When readers have automatized their word-decoding skills, they accordingly have more mental capacity (i.e., more room in their working memory (WM) which functions as a mental workspace to devote to understanding the gist and important details presented in the text (Gagné, Yekovich & Yekovich, 1993). Optimal reading performance in an L2, however, cannot be achieved solely by automatization. A control process has been argued to be necessary (e.g., Block, 1992; Carrell, Gajdusek & Wise, 1998; Hacker, 1998; Kintsch, 1998).

In most routine reading contexts, L2 readers are likely to encounter unfamiliar words, syntactic structures or topics that require them to consciously or intentionally evaluate and examine alternative sources or use context clues. Therefore, when difficulty in reading arises, regulatory or control processes, as higher-level processing, such as assessing situations and monitoring current comprehension are needed because such difficulty affects the speed and effectiveness of reading. Though this metacognitive processing may slow down reading speed, it helps increase reading achievement. According to Gagné, Yekovich and Yekovich (1993), the nature of strategies is related to the control processing component in their human information-processing model which guides and monitors information processing events.

The role of strategy use in reading comprehension has thus been a topic of discussion in the L2 reading literature. Block (1992: 320) suggests that L2 readers need to be ready to "stand back and observe themselves" when they read. Carrell, Gajdusek and Wise (1998) further pointed out that what matters may not be so much what

strategies learners use, but rather the knowledge of when, how and why a strategy is to be used. Note that when some metacognitive processes such as goal setting, planning how to achieve goals, monitoring goal attainment and revising plans are deployed automatically, they lose the significance of being part of the higher-level processing because they do not appear to be *beyond* the processing event. Rather they are part of it. Hence, the flow of control in learned, automatized skills is embedded in the skill, although the sequence has a control structure (Gagné, Yekovich & Yekovich, 1993).

In the literature, cognitive and metacognitive strategies have been regarded as closely related, postulating that metacognitive strategies have a direct impact on cognitive strategies in L2 learning, use or performance (e.g., Anderson, 2005; Bachman, 1990; Bachman & Palmer, 1996; Brown *et al.*, 1983, Chamot, 2005; Faerch & Kasper, 1983; O'Malley & Chamot, 1990; Oxford, 1990; Wenden, 1991). Cognitive strategies in turn have a direct impact on L2 performance because they are involved directly in the target language use. Recently, researchers have begun to document some evidence of the hypothetical relationship between metacognitive strategies and cognitive strategies.

Reading strategy research has also revealed how strategic readers interact with a written text and how their strategic behavior is related to effective reading comprehension (e.g., Anderson, 1991; Barnett, 1988; Bernhardt, 1986; Block, 1992; Carrell, 1984, 1989, 1991; Purpura, 1997, 1998, 1999; Salataci & Akyel, 2002). It has generally been found that successful L2 readers know how to use appropriate strategies to enhance text comprehension (Alderson, 2000; Chamot *et al.*, 1989; Yang, 2002). By contrast, poor readers generally lack effective metacognitive strategies (Alderson, 2000) and have little awareness on how to approach reading (Baker & Brown, 1984). They also have deficiencies in the use of metacognitive strategies to monitor their understanding of texts (Pitts, 1983).

Language testing research on strategic processing

Since a language test is a means to infer an individual's L2 reading ability as well as to assist decision-making in an individual's performance, test validation research is needed to inform us of possible factors that affect language test performance. As Alderson

(2000) points out, we are not interested in how well a test taker can perform a reading test, but rather we are interested in the kind of inferences we can make based on test takers' reading scores. Much of what is involved during reading test-taking should be similar to that in non-test reading. Hence, strategies such cognitive strategies (e.g., comprehending, memory and retrieval) and metacognitive strategies (e.g., planning, monitoring and evaluating) that are theoretically encapsulated in L2 reading in a non-test language use context (i.e., for making sense of the language material or task) should contribute to the reading test performance.

In the past decades, there has been research that aims to understand the nature of cognitive and metacognitive strategies that influence language test performance. Purpura (1999) examined the relationship between perceived cognitive and metacognitive strategy use and language test performance (UCLES's First Certificate in English (FCE) Anchor Test), through the applications of the SEM approach with 1,382 learners. The participants answered the context-free strategy use questionnaire prior to the test taking. Purpura found that cognitive processing was a *multi-dimensional* construct constituting a set of *comprehending, memory and retrieval* strategies. These complex cognitive strategies worked with one another to affect language performance. The model of metacognitive strategy use was a *unidimensional* construct consisting of a single set of assessment processes (e.g., goal setting, planning, monitoring, self-evaluating and self-testing). Purpura found that metacognitive processing had significant, direct and positive effects on all three components of cognitive processing (values between 0.59 and 0.86) which directly impacted the language performance.

Phakiti (2003b), through the use of a cognitive and metacognitive questionnaire drawn from the existing literature, retrospective interviews and an EFL achievement test, investigated the relationship between 384 Thai learners' cognitive and metacognitive strategy use and their reading test performance. Unlike Purpura (1999), the test-takers completed the test first and immediately after the test completion, they answered the questionnaire on the degree of their strategy use during the test taking. The rationale underlying this design was that strategy use, like other online cognitive processes would be more directly related to specific language performance than to general strategy use. Using the factor structures to form composites

of cognitive and metacognitive strategies for further quantitative analyses, Phakiti found that metacognitive strategies were statistically positively related to cognitive strategies (the correction-for-attenuation correlation = 0.76). In his qualitative data analysis, cognitive and metacognitive strategy use by successful test-takers was highly complex. For example, when they translated part of a text (cognitive strategy use), they aimed to see if it made sense (evaluating strategy use), and when they made efforts to summarize the passage (cognitive strategy use), they checked for comprehension (monitoring strategy use). In regards to the relationships between strategies and test performance, cognitive and metacognitive strategies were both positively correlated with the reading test performance.

Phakiti (2003b) also compared the differences in the strategy use and reading performance among highly successful, moderately successful and unsuccessful learners by means of factorial multivariate analysis of variance (MANOVA) and found the significant differences among these learner groups. There was strong evidence that the highly successful learners reported significantly higher use of metacognitive strategies than the moderately successful ones, who in turn reported higher use of these strategies than the unsuccessful ones. The qualitative data analysis further supported such findings, suggesting that the successful learners approached the test tasks more strategically than the less successful ones. In his subsequent study, Phakiti (2003a) reported the differences between males and females in terms of strategy use and L2 reading performance. Phakiti (2003a) found that although males and females did not differ in their reading performance and their use of cognitive strategies, males were found to report significantly higher use of metacognitive strategies than females. However, at the gender plus success level, no gender difference was found (e.g., highly successful males did not differ in terms of L2 reading performance and strategy use from their female counterparts).

Song (2004) investigated the extent to which cognitive and metacognitive strategy use accounted for Chinese test-takers' performance in the College English Test Band 4 through regression analyses. Song employed a revised strategy questionnaire mainly based on Purpura (1999). Song found that cognitive and metacognitive strategies accounted for 8.6% of the test score. In the context of the Michigan English Language Assessment Battery

(MELAB) with 161 test-takers, Song (2005) found that test-takers' perceptions of cognitive strategy use fall into six dimensions (i.e., repeating/confirming information strategies, writing strategies, practicing strategies, generating strategies, applying rules strategies and linking with prior knowledge strategies), whereas their metacognitive strategy use perceptions fall into three factors (i.e., evaluating, monitoring and assessing). The effects of strategy use on the language performance were found to be weak (explaining about 12.5 to 21.4% of the score variance).

Implications for the present study

Findings and issues raised in the literature review have implications for the design of the present study. First, cognitive and metacognitive strategies are conscious processes relevant to language test performance. Hence, to explain the nature of language performance, both cognitive and metacognitive strategies need to be taken into account. Second, cognitive and metacognitive strategies tend to contribute differently to language test performance. Based on previous research, metacognitive strategies have been found to directly regulate cognitive strategy use which in turn directly influences communicative language use. Planning strategies direct the course of individuals' thinking which indirectly results in specific cognitive strategy use for specific tasks. Planning helps allocate resources to the current task (via monitoring strategies), determine the order of steps to be taken to complete the task and set the intensity or the speed at which one should work on the task (evaluating strategies). Monitoring and evaluating strategies help identify the task on which one is currently working, check on the current progress on that task, evaluate that progress and predict what the outcome of that progress will be.

Second, in the area of statistical approaches, most previous strategy research analyzed the data in terms of the relationship of strategy use to L2 performance by means of frequency counts, correlational analysis, regression analysis, path analysis and analysis of variance (ANOVA) or t-tests. Although these statistical approaches can yield useful knowledge regarding the nature of strategy use and its relationship to L2 learning or performance, they are subject to certain analytic limitations (e.g., measurement error variance is not presumed). The concept of measurement error derives from reliability

theory (Mehrens & Lehmann, 1984) which partitions the variance of measures into true and error variance. The true score variance is related to the construct of interest, whereas the error variance is the unreliable variance that comes with the measure. The effect of unreliable variable measurements and their effect on multiple regression can be dramatic (Fuller & Hidioglou, 1978). For example, if R^2 is 0.42 and the dependent variable reliability is 0.85, only 35.7 percent (0.42×0.85) of the variance is true variance. Based on this example, if the measurement error is ignored, the relationship between the construct of interest could be estimated inaccurately because it is not related to any other measures and exists independently of other measures. Measurement error can result in either underestimating or overestimating the strength of relationships from the model that uses imperfectly measured variables in an unpredictable way. An attempt to extract the theoretical variable of interest from the unwanted variables (e.g., error and method variance; see Maruyama, 1998) has led to the application of multiple measures of a construct and the logic of factor analysis known as the linear SEM approach.

The SEM approach (as in Purpura, 1999) builds upon factor analysis, multiple regression and path analysis approaches to resolve the problem of single observed variables and their related measurement error in path analysis. It has been argued that multiple observed variables are preferred over a single variable in determining a latent variable (Pedhazur & Schmelkin, 1992) and in resolving the effects of measurement error on parameter estimates (Byrne, 1994; Kline, 1998). SEM models differ from path analysis models in that SEM models use latent variables rather than observed variables and combine a measurement model with a structural model to substantiate theory. The SEM approach thus allows researchers to focus on the construct validation of theoretical propositions (McArdele & Bell, 2000). SEM directly evaluates the effects of measurement error by considering the score reliability (i.e., $[1 - \text{measurement error variance}] / \text{total score variance}$) as part of model fitting (Stevens, 1996). In the present study, the SEM approach is used to deal with data analysis in regards to the relationships between strategy use and L2 reading variables.

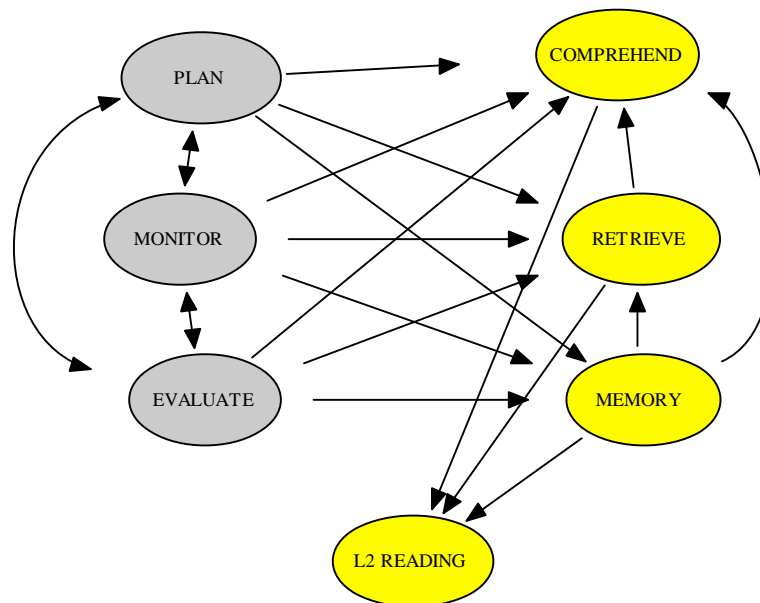
Research questions and hypotheses

The review of the literature has suggested that metacognitive strategies exert an executive role on cognitive strategies which in turn influence success in L2 performance. In the present study, there are two hypotheses to be tested. Figure 1 demonstrates plausible hypothesized relationships between cognitive strategies and metacognitive strategies and their relationship to L2 reading test performance.

Hypothesis 1: Metacognitive strategies have a direct impact on cognitive strategies.

Hypothesis 2: Cognitive strategies in turn have a direct impact on L2 reading.

Figure 1 The hypothesized relationships of cognitive and metacognitive strategies to L2 reading test performance



There are six research questions that these hypotheses could be used to answer:

Research question 1: What is the nature of cognitive strategies?

Research question 2: What are the interrelationships among cognitive strategies (comprehending, retrieval and memory strategies)?

Research question 3: What is the nature of metacognitive strategies?

Research question 4: What are the interrelationships among metacognitive strategies (planning, monitoring and evaluating strategies)?

Research question 5: What is the nature of EFL reading test performance?

Research question 6: To what extent do cognitive and metacognitive strategies affect EFL reading test performance?

3. Method

Setting and participants

A sample of 358 Thai university students (160 males and 198 females) participated in the study. The students were enrolled in a Basic English Reading and Writing subject which focuses on academic reading for undergraduate studies. They were asked to voluntarily participate in the study and were informed of the research procedures prior to the data-gathering periods. Participants identities were anonymous and their background information was confidential. For the purpose of the study, the data were gathered during their final test period (30%). The students were between the ages of 19 and 22 (mean age = 20.48; SD = 0.78) and had been studying English in Thailand for about ten years (mean year = 10.25; SD = 1.65). Their English proficiency level ranged from lower intermediate to upper intermediate. The participants took a reading comprehension test which lasted three hours. After they completed the test, they were asked to answer a questionnaire on their cognitive and metacognitive strategy use.

Measurement instruments

There were two sets of measurement instruments in this study: (1) an EFL reading comprehension test; and (2) a cognitive-metacognitive strategy use questionnaire.

EFL reading test

The test was organized around various reading tasks with two major parts. There were 85 questions in total: (1) Rational Cloze; and (2) Text Comprehension. The purposes of the two test parts differ in terms of the underlying theoretical reading constructs being measured and in terms of the nature of tasks presented.

Section 1: Rational Cloze. This section was designed to measure the readers' ability to comprehend texts using both structural and lexical appropriacy, pragmatics and discourse. The first part of this test section was multiple-choice rational cloze where test takers simply chose the most appropriate answer from the given four choices, whereas the second part was open-ended cloze where test takers needed to generate an appropriate word on their own in order to complete the text. A contextually acceptable response scoring method was used for this part.

Section 2: Text Comprehension. This section aimed to measure the readers' ability to comprehend English texts for main ideas, details and inferences. This section was composed of two sections: expeditious reading (i.e., skimming and scanning) and Careful Text Comprehension. The specific reading skills which students needed to demonstrate were: (1) scanning and skimming text for general and specific information; (2) recalling word meanings; (3) evaluating information; (4) guessing meanings of unknown words from context clues; (5) identifying the meaning of key vocabulary items in the text; (6) identifying phrases or word equivalence; (7) predicting topics of passages and the content of a passage; (8) discriminating between more or less important ideas; (9) distinguishing facts from opinions; (10) analyzing reference words in the text; (11) drawing inferences from the content; (12) identifying the title of the text and the appropriate heading; (13) summarizing the content of the given text; (14) recognizing main ideas or purposes of a passage (and distinguish them from supporting ideas); (15) synthesizing information across more than one paragraph in the text; (16) recognizing and recover

information in the form of specific details; and (17) recognizing inferences drawn from the statements and information presented in the text. This section consisted of 5 texts ranging from 150 words to 700 words. The texts and words were general/nontechnical in nature to tap into the students' language knowledge as taught in the class. For both major test sections, the topics included family, food and drink, clothing, health, travels, transportation, information technology, business and agriculture.

Cognitive and metacognitive strategy questionnaire

Methods typically used to understand the nature of strategies include verbal reports (e.g., think-aloud protocols, retrospective interviews) and self-report questionnaires. In the present study, a likert-scale questionnaire was used. In a context of a large-scale study, it can also be difficult or impossible to tape-record all participants while taking the reading test. Furthermore, the think-aloud methodology is highly complex and the participants need a lot of practice prior to actual data gathering to achieve optimal think-aloud validity. The usefulness of likert-scale questionnaires is supported by many strategy researchers (e.g., O'Malley & Chamot, 1990; Oxford, 1996; Purpura, 1999) and SEM researchers (e.g., Bentler, 1995; 2006; Byrne, 1994; Kline, 1998). The strategy questionnaire in this study was adopted from the questionnaire used by Phakiti (2003b). Phakiti (2003b) reported the construct validation of the questionnaire. Since Phakiti (2003b) identified some problematic items in his questionnaire, only 30 items that provided a clear structure of cognitive and metacognitive strategies were adopted and re-modified. Table 1 presents the taxonomy of the questionnaire.

In this study, the questionnaire was piloted for item-level analysis such as reliability estimates prior to its actual use. The questionnaire was given in Thai in order to prevent language problems in measuring their cognitive and metacognitive strategy use. The questionnaire used in this study allowed learners to mark strategy use on a 5-point Likert scale: 1 (Never), 2 (Sometimes), 3 (Often), 4 (Usually) and 5 (Always). The length of time needed to complete the questionnaire ranged from approximately 10-15 minutes. Appendix A provides the questionnaire in English.

Table 1 Taxonomy of the strategy questionnaire.

Processing	Subscale	No. of items	Items
1. Cognitive strategies	Comprehending	5	2, 3, 6, 7, 14
	Memory	4	1, 5, 8, 22
	Retrieval	4	4, 9, 26, 29
2. Metacognitive strategies	Planning	6	10, 11, 19, 20, 23, 27
	Monitoring	6	12, 16, 17, 21, 24, 25
	Evaluating	5	13, 15, 18, 28, 30
		30	

Validating the reading test

Rasch Item Response (IRT) was employed by means of the Quest Program (Adams & Khoo, 1996) to evaluate the quality of the test to be used to infer the learners' English reading achievement. It was found that the person separation reliability (equivalent to KR-20) of the total test was acceptable (0.88). The reliability estimate for Sections 1 and 2 were 0.83 and 0.86, respectively. For the purpose of this study, misfitting L2 test takers were eliminated for the SEM analysis. The mean square statistics values (Mean = 1.00; SD = 0.14) were used to determine misfitting test takers. This means that L2 test takers whose mean square statistics were 1.28 or above were misfitting ($1 \pm 2SD$). The term misfitting here is a statistical term for Rasch IRT. The fit statistics are expressed as mean square or *t* statistics which enable researchers to investigate the coherence of a test taker's responses as part of a set of responses from a larger group of test takers (McNamara, 1996). Misfitting test takers were those whose reading abilities were not measured appropriately by this particular test.

Then empirically-based, composite reading variables to be used in the SEM analyses were generated. Based on the test subsections, there were a total of four observed reading variables that were used to test as a SEM measurement model of EFL reading performance. Variables 1 (GR-Lex1) and 2 (GR-Lex2) represented the ability to comprehend English texts using both structural and lexical appropriacy, pragmatics and discourse (as derived from two sections of multiple-

choice and construct-response rational cloze). Variables 3 (TxtCOM1) and 4 (TxtCOM2) represented the ability to read English texts for main ideas, details and inferences. Variable 3 was derived from expeditious reading (i.e., performance in the skimming and scanning test section), and Variable 4 was from the careful text comprehension section. Table 2 presents the descriptive statistics for the four observed reading variables. All values for variable skewness and kurtosis were within acceptable limits (values centered near zero), which was suggestive of univariately normal distributions.

Table 2 Distributions for reading score variables

Variable	No.	Min	Max	Mean	SD	Skewness	Kurtosis
GR-LEX1	20	5.00	19.00	12.22	2.82	-0.151	-0.341
GR-LEX2	25	5.00	23.00	14.12	3.34	0.018	-0.230
TxtCOM1	15	3.00	15.00	09.47	2.06	-0.218	0.233
TxtCOM2	25	5.00	24.00	13.14	3.76	0.171	-0.663

GR-LEX = Grammatical-Lexical Reading Ability

TxtCOM = Text Comprehension Ability

Validating the strategy questionnaire

Prior to modeling the nature of cognitive and metacognitive strategy use, a number of item-level analyses were conducted to validate the constructs of cognitive and metacognitive strategies. First, the item distribution (as reported in Table 3) was examined, reliability analysis was conducted, relatively low alpha items were eliminated and subsequently a confirmatory factor analysis (CFA) for each measurement model was conducted. Cognitive strategy factors were composed of comprehending strategies, memory strategies and retrieval strategies. The metacognitive strategy factors consisted of planning strategies, monitoring strategies and evaluating strategies. The descriptive statistics for thirty observed variables are presented in Table 3. The items could be considered normally distributed. All variable skewness and kurtosis statistics were within the acceptable limits, which was suggestive of univariately normal distributions.

Table 3 Distributions for the cognitive and metacognitive strategy variables

Item	Min	Max	Mean	SD	Skewness	Kurtosis
1	1.00	5.00	2.078	1.028	0.696	-0.141
2	1.00	5.00	3.720	0.973	-0.406	-0.481
3	1.00	5.00	3.379	0.970	-0.323	-0.417
4	1.00	5.00	3.397	0.986	-0.071	-0.439
5	1.00	5.00	3.659	0.895	-0.333	-0.070
6	2.00	5.00	3.953	0.916	-0.498	-0.628
7	1.00	5.00	3.302	0.994	-0.016	-0.430
8	1.00	5.00	3.583	0.954	-0.212	-0.545
9	1.00	5.00	3.771	1.000	-0.457	-0.414
10	1.00	5.00	2.975	1.145	0.083	-0.720
11	1.00	5.00	3.106	1.076	-0.213	-0.538
12	2.00	5.00	3.763	0.821	-0.115	-0.616
13	1.00	5.00	3.383	0.938	-0.137	-0.294
14	2.00	5.00	3.855	0.847	-0.301	-0.566
15	1.00	5.00	3.737	0.874	-0.321	-0.191
16	1.00	5.00	3.670	0.945	-0.239	-0.404
17	2.00	5.00	3.913	0.898	-0.411	-0.661
18	1.00	5.00	3.539	0.880	-0.095	-0.235
19	1.00	5.00	3.285	0.877	-0.039	-0.209
20	1.00	5.00	3.240	0.822	-0.288	-0.066

Table 3 Distributions for the cognitive and metacognitive strategy variables (continued)

Item	Min	Max	Mean	SD	Skewness	Kurtosis
21	1.00	5.00	3.620	0.970	-0.471	-0.096
22	1.00	5.00	3.639	0.898	-0.110	-0.656
23	1.00	5.00	3.486	0.871	-0.072	-0.206
24	1.00	5.00	3.408	0.820	-0.008	-0.062
25	1.00	5.00	3.379	0.935	-0.122	-0.539
26	1.00	5.00	3.385	0.939	-0.084	-0.090
27	1.00	5.00	3.365	0.874	-0.229	0.045
28	1.00	5.00	3.595	0.898	-0.299	-0.243
29	1.00	5.00	3.463	0.864	-0.162	0.049
30	1.00	5.00	3.829	1.043	-0.473	-0.724

Table 4 presents the reliability estimates of the strategy composites. In summary, in this study, each participant had 3 cognitive strategy variables, 3 metacognitive strategy variables and 4 reading performance variables.

Table 4 Composites of cognitive and metacognitive strategies and EFL reading scores with internal consistency estimates

Processing	Subscale	No. of items	Items	Alpha
1. Cognitive strategies	Comprehending	5	2, 3, 6, 7, 14	0.72
	Memory	4	1, 5, 8, 22	0.51
	Retrieval	4	4, 9, 26, 29	0.63
2. Metacognitive strategies	Planning	6	10, 11, 19, 20, 23, 27	0.71
	Monitoring	6	12, 16, 17, 21, 24, 25	0.71
	Evaluating	5	13, 15, 18, 28, 30	0.68
Total		30		0.91

Establishing the structural equation models

The Statistical Packages for Social Sciences (SPSS) program version 12 was used to compute descriptive statistics and perform reliability analyses. SPSS program was used as a data file manager prior to using the EQS 6.0 program (Bentler, 1985-2006) to impute data, perform missing and outlier data analyses, confirmatory factor analyses (CFAs), covariance structure analyses and finally full-latent SEM. The SEM procedures involve (1) developing measurement models to define latent variables and then (2) establishing relationships among the latent variables.

The EQS refers to observed variables as V (typically diagrammed using rectangles or squares) and to latent or unobserved variables as F (typically diagrammed using ellipses, ovals or circles). A residual associated with a V is labeled E to denote specificity and measurement error, whereas a residual associated with a F is labeled D (disturbance) to represent error in the prediction of one factor from another. A one-way arrow represents a standardized regression coefficients or factor loading, whereas a curved two-way arrow represents a covariance or correlations between variables. In a standardized solution, all V, F, E and D variables are rescaled to have a variance of 1.0. The standardized solution makes it easy to interpret the variables in the linear structural equation system (see Bentler, 2006). One feature of the standardized solution is that previously fixed parameters (e.g., 1.0) will take on new values. Note that the standardized solution produced by the EQS program is not the same solution as provided by Jöreskog and Sörbom (1988), who do not standardize measured variables, errors in variables or disturbances in equations (Bentler, 2006; Bryne, 1994).

Each measurement model was tested before the structural relationships were finally tested simultaneously because model misfit in the full latent SEM could initially derive from the misspecification at the level of measurement models. Once each measurement model was specified, its plausibility based on sample data comprising all observed variables in the model was tested. In the present study, the maximum likelihood (ML) estimation method in model estimations was used because the observed variables were ordinally-scaled and multivariately-normal. ML estimation is typically used to seek parameters that best reproduce the estimated population variance-

covariance matrix. The evaluation of model adequacy is usually based on an inspection of the values of standardized residuals, the chi-square statistics, other fit indices and researchers' knowledge of the data and theoretical and conceptual aspects of the constructs under study (see e.g., Bentler, 1995; Byrne, 1994; Schumacker & Lomax, 1996, for an extensive discussion of these criteria). Some observed variables originally thought to represent the constructs of interest were excluded from some measurement models. The items to be excluded were Items 1, 3, 10, 11, 25 and 30.

After establishing the measurement models, the direct influence of the metacognitive strategies on the cognitive strategies was tested. To achieve this, a path coefficient regressing from metacognitive strategies to cognitive strategies (i.e., cognitive strategies were regressed on metacognitive strategies) was estimated. Then a path coefficient regressing from cognitive strategies to EFL reading test performance estimated. It was found that the original hypothesized model as presented in Figure 1 did not fit well with the data and hence a number of competing SEM models were tested and retested (see Bentler, 2006, for model respecification). In this study, errors among some cognitive and metacognitive strategy use variables were corrected (see Figure 2). Following Bentler's (2006) advice, the shared error variances due to the contents of the measures could be corrected by correlating the errors, so that the model was better explained.

4. Results and discussion

Figure 2 presents the full-latent SEM model that best represents the data in the present study. The independence chi-square statistic ($\chi^2_{(378)}$) was 3847.619. The Chi-square statistic of the hypothesized SEM model ($\chi^2_{(330)}$) was 612.470. The large difference in the chi-square values between the independence and tested models suggests that the tested model has good model fit¹. The probability value for the chi-square statistic of the hypothesized model was significant ($p = 0.000$). Unlike other standard statistical analyses, SEM researchers need to

¹ Given a sound hypothesized model, the χ^2 value for the null model needs to be high, thereby suggestive of excessive malfit. The chi-square of the alternative model, on the other hand, needs to be relatively much lower than that of the null model.

obtain a nonsignificant χ^2 ($p > 0.001$). The fit indices consistently indicated a good model fit (e.g., Bentler-Bonett Non-normed Fit Index = 0.91; CFI = 0.92; RMSEA = 0.049)². An examination of the appropriateness or feasibility of parameter estimates and the statistical significance of parameter estimates indicated that all estimates were reasonable and statistically significant at the 0.05 level.

Table 5 presents the standardized solution of this model, along with the value of R^2 . One appropriate category of Magnitude-of-Effect (ME) estimates which is related to the SEM approach is the use of R^2 (Stevens, 1992), given that the units of measurement are to be meaningful on a practical level. This ME measure includes an index that involves proportions of variance (i.e., how much of the variability in the dependent variable(s) is associated with the variation in the independent variable(s)). The magnitudes range from 0 to 1. For the purpose of communication, observed items (see Table 5) were given names based on the taxonomy. For example, item 2 was named as 'translate'.

I now turn to discuss answers relevant to the research questions and the hypotheses.

RQ 1 (What is the nature of cognitive strategies?)

Figure 2 shows that cognitive strategies are composed of comprehending strategies, retrieval strategies and memory strategies. *Comprehending strategies* were explained by V2 (translating) with a loading of 0.51 ($R^2 = 0.26$), V6 (focus on meaning) with a loading of 0.61 ($R^2 = 0.37$), V7 (scanning and skimming) with a loading of 0.57 ($R^2 = 0.32$) and V14 (main points) with a loading of 0.74 ($R^2 = 0.55$).

² Theoretically, if a Comparative Fit Index (CFI), for example, equals 0.95, the relative overall fit of the researcher's model is 95 percent better than that of the null model estimated with the same sample data. The fit indices of 0.90 suggest good model fit (Bentler, 1995). For the standardized root mean squared residual (SRMR) which is a standardized summary of the average covariance residuals - the differences between the observed and model-implied covariances and the root mean squared error of approximation (RMSEA), values of 0.05 or less is desired because when the fit of the model is perfect, these statistics equals zero.

Based on the computation of correlation coefficients among these observed strategies, the correlation coefficients³ ranged from 0.29 (V2 and V7) to 0.45 (V6 and V14).

Based on Schumacker and Lomax (1996), the *total* common factor variance (h^2) that explains how much of the theoretical construct is explained by the shared common variance in the set of variables loading on the single factor) was 0.37⁴. This indicated that only 37 percent of the comprehending strategy factor variance was defined by the four variables. Based on this, the unique (residual) factor variance was accounted for 63 percent. It should be noted that for a unidimensional indicator, standardized factor loadings should have absolute values less than 1.00 because they are correlations (Kline, 1998). This result demonstrates that although the latent variable has the same scale as one of its indicators, it is not identical to that indicator. Hence, it is almost impossible that the total common factor variance could explain 100 percent of the latent factor variance. In summary, the nature of comprehending strategies are related to focusing on meaning, translating, identifying main points in the text and skimming and scanning the text for effective comprehension.

³ The correlation coefficients among observed variables could be simply computed by multiplying the weights or factor loadings between pairs of variables. For example, V2 and V6 had a correlation coefficient of 0.31 (i.e., 0.51×0.61).

⁴ Based on Schumacker and Lomax (1996), h^2 can be calculated by squaring each of the weights (factor loadings), summing them up and then dividing the sum by the number of variances. Hence, h^2 of cognitive strategy use is $(0.51^2 + 0.61^2 + 0.57^2 + 0.74^2)/4 = 0.37$. The unique (residual) factor variance can be computed as: $1 - h^2$.

Table 5 Standardized parameter estimates for cognitive and metacognitive strategies and EFL reading performance

Observed Variable	Label	Loading	Error	R-squared (R ²)
Item 2	V2	Translate	0.505*F4 + 0.863 E2	0.255
Item 4	V4	Use grammar	0.486* F5 + 0.874 E4	0.236
Item 5	V5	Devote time	0.553* F6 + 0.833 E5	0.306
Item 8	V6	Focus on meaning	0.610* F4 + 0.793 E5	0.372
Item 7	V7	Scan & Skim	0.569*F4 + 0.822 E7	0.324
Item 8	V8	Repetition	0.523*F6 + 0.852 E8	0.273
Item 9	V9	Prior knowledge	0.563*F5 + 0.827 E9	0.317
Item 12	V12	Aware of what and how	0.554*F2 + 0.832 E12	0.307
Item 13	V13	Check performance	0.543*F3 + 0.840 E13	0.295
Item 14	V14	Main points	0.741*F4 + 0.672 E14	0.548
Item 15	V15	Assess meaning	0.731*F3 + 0.682 E15	0.535
Item 16	V16	Aware of strategy	0.574*F2 + 0.819 E16	0.329
Item 17	V17	Correct mistakes	0.549*F2 + 0.836 E17	0.301
Item 18	V18	Self-test	0.533*F3 + 0.846 E18	0.284
Item 19	V19	Set task expectation	0.662*F1 + 0.750 E19	0.438
Item 20	V20	Plan actions	0.631*F1 + 0.776 E20	0.398
Item 21	V21	How much to do	0.497*F2 + 0.868 E21	0.247
Item 22	V22	Remember tasks	0.674*F6 + 0.739 E22	0.454

Table 5 Standardized parameter estimates for cognitive and metacognitive strategies and EFL reading performance (continued)

Observed Variable	Label		Loading	Error	R-squared (R ²)
Item 23	V23	Set goals	0.698*F1	+ 0.716 E23	0.487
Item 24	V24	Monitor processes	0.656*F2	+ 0.754 E24	0.431
Item 26	V26	Multi-strategies	0.517*F5	+ 0.856 E26	0.268
Item 27	V27	Clarify goals	0.619*F1	+ 0.785 E27	0.383
Item 28	V28	Check accuracy	0.558*F3	+ 0.830 E28	0.312
Item 29	V29	Connect relevant information	0.636*F5	+ 0.772 E29	0.404
GR-LeX1	V31	GR-LeX1	0.791*F7	+ 0.612 E31	0.626
GR-LeX2	V32	GR-LeX2	0.815*F7	+ 0.580 E32	0.664
TxtCOM1	V33	TxtCOM1	0.612*F7	+ 0.791 E33	0.375
TxtCOM2	V34	TxtCOM2	0.683*F7	+ 0.730 E34	0.467
Comprehend	F4	Comprehend	0.936 *F5	+ 0.352 D4	0.876
Retrieval	F5	Retrieval	0.402 *F6 + 0.620 *F3	+ 0.000 D5	1.000
Memory	F6	Memory	0.985 *F2	+ 0.286 D6	0.918
EFL Reading	F7	EFL Reading	0.573 *F4	+ 0.820 D7	0.328

F1 = Planning Strategies**F2** = Monitoring Strategies**F3** = Evaluating Strategies**F4** = Comprehending Strategies**F5** = Retrieval Strategies**F6** = Memory Strategies**F7** = EFL reading performance

Retrieval strategies were explained by V4 (use grammar) with a loading of 0.49 ($R^2 = 0.24$), V9 (use prior knowledge) with a loading of 0.51 ($R^2 = 0.26$), V7 (multiple strategies) with a loading of 0.52 ($R^2 = 0.27$) and V29 (connect relevant information) with a loading of 0.64 ($R^2 = 0.41$). Based on the computation of correlation coefficients among these observed strategies, the correlation coefficients ranged from 0.25 (V4 and V26) to 0.36 (V9 and V29). The *total* common factor variance (h^2) was 0.44. This indicated that 44 percent of the retrieval strategy factor variance was defined by the four variables. Based on this, the unique (residual) factor variance accounted for 65 percent of the total variance. In summary, retrieval strategies are related to making use of grammatical knowledge, prior knowledge, multiple strategies and relevant information.

Memory strategies were explained by V5 (devote time to experienced difficulty) with a loading of 0.55 ($R^2 = 0.30$), V8 (repetition) with a loading of 0.52 ($R^2 = 0.27$), and V22 (remember tasks) with a loading of 0.67 ($R^2 = 0.45$). Based on the computation of correlation coefficients among these observed strategies, the correlation coefficients ranged from 0.14 (V5 and V8) to 0.37 (V5 and V22). The *total* common factor variance (h^2) was 0.34. This indicated that only 34 percent of the comprehending strategy factor variance was defined by the four variables. Based on this, the unique (residual) factor variance accounted for 64 percent of the total variance. In summary, memory strategies are related to memorizing information and tasks such as via time devotion, repetition and understanding questions adequately before completing the test task.

The structures of these cognitive strategies were found to be similar to those found by Purpura (1999). Nevertheless, the supposed roles of these cognitive strategies for this reading context could be far more complex than what this SEM model suggests because indeed cognitive strategies are multidimensional as it can be seen as (1) general, (2) domain-specific and (3) task-specific. General cognitive strategies have wide utility across different content areas and cognitive tasks, for example, note-taking, rehearsal, elaboration, underlining, mapping and networking and summarization. Domain-specific cognitive strategies can be seen as closely related to use of the target language (e.g., translation, guessing meanings of unknown words and applying grammar rules) and world knowledge to perform the given L2 reading tasks. Task-specific cognitive strategies

are related to certain kinds of reading tasks. For example, task-specific cognitive strategies associated with a multiple-choice reading test include elimination of bad alternatives, recognizing parallelism, looking for the portion of the text that the question refers to and then looking for clues to the answer and looking for answers to questions in chronological order in the text, etc (see Cohen, 1998, in press). Some of these cognitive strategies were missing in this hypothesized model. However, one implication suggested from the nature of the factor loadings of cognitive strategies found in this study may be that in this L2 context of reading comprehension, comprehending, retrieval and memory strategies were conscious behaviors relevant to EFL reading as they helped the reader make sense of the language materials.

RQ2 (What are the interrelationships among cognitive strategies [comprehending, retrieval and memory strategies]?)

Based on Figure 2, it was found that the relationships among comprehending strategies, retrieval strategies and memory strategies were uni-directional. That is, contrary to what was originally thought, the relationships were not reciprocal. First, based on the testing and retesting of the hypothesized SEM model, it was found that memory strategies (via the control of monitoring strategies; as discussed below) influenced the extent to which retrieval strategies (via the control of evaluating strategies; as discussed below) were to be used. The regression coefficient of memory strategies on retrieval strategies was 0.40 ($R^2 = 0.16$). Second, through retrieval strategies, memory strategies were found to indirectly affect comprehending strategies (regression coefficient = 0.38; $R^2 = 0.14$). Third, retrieval strategies were found to largely affect comprehending strategies (regression coefficient = 0.94; $R^2 = 0.88$). This means that comprehending strategies depend largely on the effectiveness of retrieval strategies.

Perhaps in reading comprehension, the extent to which comprehending strategies are successfully used depends on how well one can retrieve information (e.g., long-term linguistic knowledge and world knowledge, and newly stored information in the current text) from the memory. In other words, it can be argued that comprehending strategies would be the result of integration between memory and retrieval strategies. This finding is interesting in that in test taking contexts, the ability to retrieve information effectively

seems to positively result in better use of comprehending strategies. However, in this kind of context, memory strategies may not contribute much to comprehending strategies. Purpura (1999) pointed out similar findings in that in a test context, successful test takers may be *product-oriented* rather than process-oriented because attention is given to retrieving information, instead of memorizing it. That is, unlike in a learning context where learners need to store the language in their memory, in a test taking context, memorizing may add extra constraints to the flow of cognitive processing, thereby slowing down the process of comprehending text. Nevertheless, in this study, the relationships among these cognitive strategies make sense as readers need to memorize information and retrieve it for comprehension and performance demonstration purposes.

RQ 3 (What is the nature of metacognitive strategies?)

Figure 2 shows that metacognitive strategies are composed of planning strategies, monitoring strategies and evaluating strategies. *Planning strategies* were explained by V19 (set task expectations) with a loading of 0.66 ($R^2 = 0.44$), V20 (plan actions) with a loading of 0.63 ($R^2 = 0.40$), V23 (set goals) with a loading of 0.70 ($R^2 = 0.49$) and V27 (clarify goals) with a loading of 0.62 ($R^2 = 0.38$). The correlation coefficients ranged from 0.25 (V20 and V27) to 0.46 (V19 and V23). The h^2 was 0.43 and the true unique (residual) factor variance was 0.67. In summary, planning strategies are related to the conscious behaviors of goal setting and planning to achieve the desired goals.

Monitoring strategies were explained by V12 (awareness of what and how) with a loading of 0.55 ($R^2 = 0.30$), V16 (awareness of strategies) with a loading of 0.57 ($R^2 = 0.32$), V17 (correct mistakes) with a loading of 0.55 ($R^2 = 0.30$), V21 (how much to do) with a loading of 0.50 ($R^2 = 0.25$) and V24 (monitor processes) with a loading of 0.66 ($R^2 = 0.44$). The correlation coefficients ranged from 0.28 (V12 and V21; V17 and V21) to 0.38 (V16 and V24). The h^2 was 0.32 and the true unique (residual) factor variance was 0.68. In summary, monitoring strategies are related to awareness of what and how one is doing, mistake correction, awareness of strategies being used, how much has been achieved and to be achieved and monitoring of information processing.

Evaluating strategies were explained by V13 (check performance) with a loading of 0.54 ($R^2 = 0.29$), V15 (assess meaning) with a loading of 0.73 ($R^2 = 0.53$), V18 (self-test) with a loading of 0.53 ($R^2 = 0.28$) and V28 (check accuracy) with a loading of 0.56 ($R^2 = 0.31$). The correlation coefficients ranged from 0.29 (V13 and V18) to 0.41 (V15 and V28). The h^2 was 0.36 and the true unique (residual) factor variance was 0.64. In summary, evaluating strategies are related to conscious processing of assessing accuracy of performance and comprehension by self-testing.

RQ4 (What are the interrelationships among metacognitive strategies [planning, monitoring and evaluating strategies]?)

Based on Figure 2, it was found that the relationships among planning strategies, monitoring strategies and evaluating strategies were multi-directional. These strategies were found to work together closely to affect cognitive strategies (as further discussed in RQ6). It was found that the correlation coefficient between planning and monitoring was 0.89 ($R^2 = 0.79$). The correlation coefficient between planning and evaluating was 0.84 ($R^2 = 0.71$). The correlation coefficient between monitoring and evaluating was 0.95 ($R^2 = 0.90$). These coefficients were reasonable and supported by previous research (e.g., Phakiti, 2003b; Purpura, 1999). Particularly, monitoring and evaluating tended to work together to influence cognitive strategies (as discussed below). It can be argued that L2 test takers need to be capable of employing the following kinds of strategies effectively: (1) planning strategies to determine how to complete and solve reading task difficulty; (2) monitoring strategies to control ongoing performance; and (3) evaluating strategies to judge or make decisions about current performance. These strategies are strongly associated with the ability to assess situations of language use (Bachman & Palmer, 1996). The significant relationships among planning, monitoring and evaluating strategies provide empirical evidence in support of the notion put forth by a number of researchers (e.g., Bachman & Palmer, 1996; Purpura, 1999) that these metacognitive strategies are highly interrelated.

Nevertheless, the interrelationships between metacognitive strategies could be even more complex than the one found in this study. In the context of a reading test, metacognitive strategy use would also be about calling upon test-management and test-wiseness strategies

associated with a particular format (e.g., gap-filling or comprehension tasks) and monitoring individuals' affective states such as test anxiety or emotion and deciding how to deal with negative emotional states. Hence, in a context of this sort, at any moment in the reading and test-taking processes, the flows of metacognitive strategies could fluctuate from dealing with the test and time management to dealing with specific difficulties arising from particular reading texts or tasks. Given this possible complexity, a comprehensive online strategy use assessment during their test completion would yield a better understanding of this strategic processing network. However, in real life, official test taking (often high-stakes), this measurement method is rarely possible since answering strategy use questionnaires at the same time of taking the test would not only interfere their test performance, but also would raise ethical considerations in research conducts.

RQ 5 (What is the nature of L2 reading performance?)

Based on the full-latent model, L2 reading performance was explained by V31 (Grammatical-Lexical 1 -- with multiple-choice options) with a loading of 0.79 ($R^2 = 0.63$), V32 (Grammatical-Lexical 1 - with constructed response) with a loading of 0.81 ($R^2 = 0.66$), V33 (text comprehension 1 - expeditious reading) with a loading of 0.61 ($R^2 = 0.37$) and V34 (text comprehension 2 - careful reading) with a loading of 0.68 ($R^2 = 0.46$). The correlation coefficients between the observed variables ranged from 0.41 for V33 and V34 to 0.64 for V31 and V32. The h^2 was 0.53. The unique (residual) factor variance was 0.47. Based on the proportion between the total factor variance and the residual variance, the reading scores were a good indicator of the EFL reading ability being measured.

Unlike most studies discussed in Bernhardt and Kamil (1995), this study found that over 50 percent of the variance in L2 reading ability was explained by the factor considered. For this particular group of learners, the scores from the two rationale cloze sections seemed to be a better indicator of their EFL reading ability than from the text comprehension sections (see observed factor loadings in Table 5 and Figure 2). This noticeable result might confirm Purpura's (1999) proposition that L2 reading ability at the levels of advanced beginners and intermediates was related more to the ability to decode input text at *lexical*, *syntactic*, *semantic* and *discourse* levels than on the

ability to perform reading at higher levels of processing such as understanding main ideas, drawing inferences, using context clues and appropriate prior knowledge and synthesizing information. This position is defensible because these L2 learners might have primarily lacked specific linguistic competences such as lexico-grammatical knowledge.

RQ 6 (To what extent do cognitive and metacognitive strategies affect EFL reading test performance?)

Effects of metacognitive strategies on cognitive strategies

Based on Figure 2, it was found that planning, monitoring and evaluating strategies had differential relationships to comprehending, retrieval and memory strategies. Prior to this current SEM model, a number of alternative models that postulated the direct effects of metacognitive strategies on cognitive strategies were tested. However, it was found that not all metacognitive strategies affected cognitive strategies in a similar way. The depiction of the roles of certain metacognitive strategies on certain cognitive strategies is interesting.

First, it was found that planning strategies did not directly affect the use of comprehending, retrieval or memory strategies. However, via monitoring strategies, planning strategies were found to indirectly affect memory strategies (regression coefficient = 0.85; $R^2 = 0.72$). Based on this relationship, it may be inferred that planning tends to indirectly enhance storage of information, rather than to retrieval or comprehension of text. Via monitoring strategies and memory strategies, planning strategies indirectly affected retrieval strategies (regression coefficient = 0.34; $R^2 = 0.12$). Via evaluating strategies, planning strategies indirectly affected retrieval strategies to a greater degree (regression coefficient = 0.52; $R^2 = 0.27$). Of all the three cognitive strategies, comprehending strategies had the 'most distant' relationship to planning strategies. Via monitoring, memory and retrieval strategies, the coefficient between planning and comprehending strategies was 0.32 ($R^2 = 0.10$). Through evaluating and retrieval strategies, the coefficient between planning and comprehending strategies was 0.49 ($R^2 = 0.24$). Based on the current findings, it may be inferred that planning strategies are essential for language test performance, although they cannot be seen as directly

affecting any specific cognitive strategies because planning does not perform an executive function of assessing ongoing processes (in contrast to monitoring and evaluating strategies).

Second, monitoring strategies were found to directly affect memory strategies (regression coefficient = 0.96; $R^2 = 0.92$). This finding is interesting in that in the process of reading and test-taking, monitoring is critical to the accuracy of information storage. Without accurate monitoring or activation of monitoring strategies, memory strategies may not be effectively used. Through memory strategies, monitoring strategies had an indirect effect on retrieval strategies (regression coefficient = 0.38; $R^2 = 0.14$) and through memory and retrieval strategies, monitoring strategies had an indirect effect on comprehending strategies (regression coefficient = 0.36; $R^2 = 0.13$). In a similar way, via evaluating strategies, the regression coefficient between monitoring and retrieval strategies was 0.59 ($R^2 = 0.35$) and the regression coefficient between monitoring and comprehending strategies, through evaluating and retrieval strategies, was 0.55 ($R^2 = 0.30$). Based on the current findings, it may be argued that monitoring plays a differential role on certain cognitive strategies. Perhaps in information processing, monitoring strategies are more related to conscious behaviors directed towards storing information than retrieving it.

Third, evaluating strategies were found to directly affect retrieval strategies (regression coefficient = 0.62; $R^2 = 0.38$). Based on this finding, in retrieving information for comprehension, evaluation is found to be critical to success. Evaluating strategies were found to indirectly affect comprehending strategies (regression coefficient = 0.58; $R^2 = 0.34$) via retrieval strategies. Perhaps in reading test taking, evaluating strategies are related to conscious behaviors directed to retrieving information. Based on the present findings, it can be argued that monitoring and evaluating strategies are complementary to each other.

In summary, the present study found that metacognitive strategies explained from 10% to 92% of the cognitive strategies variance. The range of variances suggests that different metacognitive strategies function differently with other cognitive processing. It should be noted that during information processing, cognitive strategies such as memory, retrieval and comprehending may be operating

independently of any metacognitive strategies (Rumelhart, 1977) and hence their independent operation would decrease the strength of their relationship to metacognitive strategies as evident in this study.

The present findings also suggest that the flows of cognitive strategies were regulated by planning, monitoring and evaluating strategies. During information processing in different language use context, the relationship between cognitive and metacognitive strategies remains highly complex because at any moment in the reading and test-taking process, the readers may be engaged in a synchronic situation-related variation (Tarone, 1998) between cognitive strategies and metacognitive strategies. One moment they are actually checking their prior knowledge (cognitive) and the next they are asking themselves whether they have enough knowledge or information to answer the question (self-testing). Although this process would appear to be a back-and-forth relationship, it can be much more complex than that because test takers are not dealing with one strategy at a time but a sequence of strategies for one item task and a cluster of other strategies for other item tasks. The nature of such strategic processing would be parallel rather than serial (Anderson, 1995, 2000). The SEM model empirically supports the proposition that metacognitive strategies act as an executive function to regulate cognitive strategies (e.g., Bachman, 1990; Bachman & Palmer, 1996; Brown *et al.*, 1983; Gagné, Yekovich & Yekovich, 1993; Phakiti, 2003b; Purpura, 1999).

Effects of cognitive strategies on EFL reading test performance

Prior to the SEM model in Figure 2, a number of regression paths from the three cognitive strategies on EFL reading test performance were tested and retested. However, based on the statistical evaluation of model fitness, it was found that not all cognitive strategies affected EFL reading test performance in a similar manner. Based on Figure 2, it was found that memory strategies (in association with monitoring strategies) affected EFL reading test performance via retrieval and comprehending strategies (regression coefficient = 0.21 (i.e., $0.40 \times 0.94 \times 0.57$; $R^2 = 0.04$). Retrieval strategies also indirectly affected EFL reading test performance via comprehending strategies (regression coefficient = 0.53 (i.e., 0.94×0.57 ; $R^2 = 0.28$). Comprehending strategies were the only strategies that had a direct effect on EFL reading test performance (regression coefficient = 0.57; $R^2 = 0.32$).

Although memory strategies and retrieval strategies were only indirect to the test performance, they formed the foundations of comprehending strategies.

Furthermore, all metacognitive strategies⁵ were found to have indirect positive effects on EFL reading comprehension test performance. For example, monitoring strategies indirectly explained 4% of the EFL reading test performance (i.e., $0.96 \times 0.40 \times 0.94 \times 0.57 = 0.21$), whereas evaluating strategies indirectly explained 11% of the EFL reading test performance (i.e., $0.62 \times 0.94 \times 0.57 = 0.33$). Planning strategies affected EFL reading test performance through monitoring and evaluating strategies. The finding that there is an indirect relationship of the metacognitive strategies on EFL reading test performance provides a contribution to the theory of human-information processing in terms of hierarchical roles of the cognitive mechanisms (e.g., Anderson, 2000; Gagné, Yekovich & Yekovich, 1993; Kintsch, 1998).

Since the influence of cognitive and metacognitive strategies on EFL reading test performance was examined through variance and covariance analyses of individual differences data, the positive relationship between cognitive and metacognitive strategy use and EFL reading test performance suggest that successful and unsuccessful test takers differed in their cognitive and metacognitive strategy use (i.e., successful test takers reported higher degrees of cognitive and metacognitive strategy use than unsuccessful test takers). Differences in their strategy use then reflected the extent to which cognitive and metacognitive strategy use is related to the test score variance.

In summary, based on the findings in the present study, it can be argued that comprehending strategies alone do not affect EFL reading test performance, but that the simultaneous orchestration of planning, monitoring, evaluating, memory, retrieval and

⁵ A model of the direct influence of metacognitive strategies on EFL reading test performance were tested. However, the model was not explained well with the data (e.g., a number of malfitting parameters, CFI was below 0.70, etc.). A series of *post hoc* fitting suggested a misspecification with the path coefficient of metacognitive strategies on EFL reading performance.

comprehending strategies does. Although the relationship of cognitive and metacognitive strategy use to EFL reading test performance was weak, this degree of variance is meaningful within the model of human-information processing (Gagné, Yekovich & Yekovich, 1993) and communicative language ability (Bachman & Palmer, 1996).

Optimal effects of strategy use on L2 performance

According to the theory of human-information processing, cognitive processing components range from being highly conscious to highly unconscious. In the course of processing information, when routine or skilful behaviors (i.e., common processes) are not sufficient to achieve desired performance, conscious and intentional processing takes over. Among individuals, consciousness about cognitive and metacognitive strategy use may also vary. Some strategies that were identified in the questionnaire might have become automatic for some individuals, but have remained strategic for others. For example, highly proficient readers encountering an easy reading task might report low use (e.g., rare use) of a *monitoring comprehension* strategy while performing the task. Lower proficiency readers, in contrast, might also report low use (e.g., rare use) of this strategy, perhaps not because it had become automatic, but perhaps because they might not simply engage such processes in their reading and test taking. Let us assume that the former group obtains high scores while the latter one obtains low scores on the same task. The relationship between the monitoring strategy use and the reading performance among the former group would seem weaker than that of the latter group. Therefore, based on the human-information processing theory (e.g., Anderson, 2000; Gagné, Yekovich & Yekovich, 1993; Kintsch, 1998), the present findings in the relationship between strategy use and EFL reading test performance are meaningful because strategies are not always called upon to achieve the goals.

Furthermore, based on Bachman and Palmer (1996), language test performance can be understood as influenced by (1) communicative language ability (CLA) which includes language knowledge, strategic competence and affect, (2) test method effects, (3) personal characteristics and (4) random error of measurement. According to this framework, test method facets, personal characteristics and random error of measurement also explained EFL reading test

performance. Language knowledge plays a crucial role in the language test achievement because even if readers are aware of their thinking or performance, they may not be able to achieve desired performance when they lack necessary linguistic knowledge relevant to that task.

Nevertheless, knowing that strategy use is a source of good performance is significant to language assessment theory and hence suggests the importance of strategic competence in L2 use, learning or testing (Bachman & Palmer, 1996). It can be argued that success in using strategies may depend on certain conditions including: (1) when the strategy relates well to the L2 task at hand; (2) when the strategy is linked well with other strategies and processes relevant to the given task; and (3) when the strategy coordinates well with the individual's learning and cognitive style (Anderson, 2005).

5. Conclusions

The present study investigated the nature of cognitive and metacognitive strategies and their effects on EFL reading test performance. It was found that the degree of relationship between strategies varied depending on the function of cognitive processing. The nature of cognitive and metacognitive strategies and their relation to a given L2 reading test performance could, however, be far more complex than has been detected in this study due to the number of interactions among the constructs, other unmeasured constructs (e.g., automaticity) and the context in which strategy and language use occurred.

It is commonly known that researching human minds is notoriously difficult. We know that there is no ideal research method or approach to assess and analyze strategy use and L2 ability. Second, in conducting research, we have to find a balance between using a perceived ideal method and sensibility of research practice such as ethical and practical considerations. Hence, the outcomes of any empirical research analyses are merely based on inferences that do not necessarily represent actual reality of human minds. In the case of the structural equation modeling approach in particular, although structural equation modeling is enmeshed in this current knowledge and shape and guide much of what is done in this study, variance-

based modeling is not necessarily a window of any particular mind. That is, given people vary among each other, any structural model that is based on analysis of variance and covariance among people does not represent actual cognitive reality. Yet, without an effort to understand human cognitive reality via some kind of scientific analysis of empirical data, L2 theory explaining the influence of non-linguistic factors cannot be advanced.

Limitations and recommendations

There are some limitations that provide implications for further research. First, as is measuring L2 communicative language ability, measuring strategic competence is a challenging task and to date, we are faced with the problem of connecting our measures with the constructs. In this study, many inferences on the nature of cognitive and metacognitive strategies are drawn through the use of a Likert-scale questionnaire, which, of course, is limited in its capacity to capture the construct of interest. Since strategic processing is driven by mental processes that do not often lend themselves to direct observation, consequently we are heavily reliant on individuals' own accounts (Tseng, Dörnyei & Schmitt, 2006). Since none of the measurement methods is ideal (see Chamot, 2005; Wigglesworth, 2005, for discussion of elicitation techniques), we only do the best for what we can get. The decision to use a particular research instrument depends on the context of test or language use, research budget, time frames and the purpose of a study.

Second, theoretically, there is a need to simultaneously examine both perceived knowledge of how one generally uses strategies that are free from context (i.e., strategic knowledge in the long-term memory; e.g., Purpura, 1999) and perceived strategy use in an actual, specific context (i.e., online strategic regulation; as in this study). Phakiti (2003a) argued that although general perceived strategy use free from context is likely to be related to strategy use in a specific context, both pertain to different kinds of cognitive-reference classes. That is, as Kintsch (1998) pointed out, just because individuals know something about their strategy use in general does not mean that this knowledge is activated in a given cognitive process at a given time. Based on Phakiti (2003a), if empirical research shows patterns of similarity or difference between the extent to which the typical learner perceives using certain strategies over a variety of contexts and the extent to

which he/she actually uses them in a specific context, this pattern will represent something about the psychology of learners' strategic competence. The present study is limited to the fact that it does not take general perceived strategies use (i.e., metacognitive knowledge) into consideration in explaining EFL reading test performance.

The *state-trait* distinction may be a key to open door of the complex strategic processing (Phakiti, 2003b). States and traits refer to two different classes of individuals' psychological attributes. Each individual has (1) a relatively stable trait and (2) a transitory state (Spielberger, 1972). To contextualize these notions into L2 strategies, traits are strategic attributes of the L2 learners that are relatively stable across occasions (despite considerable variation in the range of settings and circumstances; as in Purpura's (1999) strategy questionnaires; e.g., '*when I read an English text, I attempt to relate the topic to my prior knowledge or experience*'). Unlike traits, *states* represent dimensions of *intra-individual* variability on different occasions. State strategies are transitory conscious processes at a given time and at a particular level of intensity (as in this study). An example of state strategy use is '*when I read this text and tried to complete the given task, I attempted to relate the topic to my prior knowledge or experience.*' The intensity of such strategy use will hence change from one occasion to another. Therefore, given the characteristics of state and trait strategy use, it is essential that they be distinguished, measured and treated in a different manner. State and trait strategies, which are the different facets of the same strategic competence, may contribute differently to language learning, acquisition and use.

Third, the findings in the present study were based on a cross-sectional design. L2 test performance has long been addressed as being highly complex, multidimensional and variable according to a variety of social and contextual factors. This complexity cannot be captured in a single performance (Spolsky, 1995). Hence, future research, with the state and trait distinction and data gathering, needs to observe *consistency* in such findings over time across various settings because instances of reading performance can be only unique to a particular context. Future studies can be conducted as a *sequential* design where different variables can be measured on successive occasions and where the model specifies effects of variables on a given occasion on other variables at later occasions. The focus of this design should be on the pattern of influences operating over time

among different variables. In a *repeated-measures* design, the same variables are measured on different occasions (e.g., a three-months interval). This design observes the relationships among repeated variables and the patterns of changes over time.

In conclusion, understanding factors affecting language test scores is crucial because we need to be able to describe and explain *variations* in language test performance and the correspondence between test performance and non-test language use (Bachman & Palmer, 1996). If a language test is to be used in social contexts, there will always be the need for an informed inference about a language test score that can be used to make a decision on an individual test-taker. Given this reasoning, a model explaining the role of non-linguistic factors, such as strategic competence, constituting *ability for use* (Hymes, 1972) is needed.

References

- Adams, R. J. & Khoo, S-K. (1996). *Quest* (Version 2.1) [Computer software]. Victoria, Australia: Australian Council for Educational Research.
- Alderson, J.C. (2000). *Assessing reading*. Cambridge: Cambridge University Press.
- Anderson, J.R. (1995). *Learning and memory: An integrated approach*. New York: John Willey & Sons.
- Anderson, J.R. (2000). *Learning and memory: An integrated approach* (2nd ed.). New York: John Willey & Sons.
- Anderson, N.J. (1991). Individual differences in strategy use in second language reading and testing. *Modern Language Journal*, 75, 460-472.
- Anderson, N. J. (2005). L2 learning strategies. In E. Hinkel. (Ed.) *Handbook of research in second language teaching and learning* (pp. 757-771). Mahwah, NJ: Lawrence Erlbaum.
- Bachman, L.F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Bachman, L.F., & Palmer, A.S. (1996). *Language testing in practice*. Oxford: Oxford University Press.

-
- Baker, L., & Brown, A.L. (1984). Metacognitive skills and reading. In P.D. Pearson (Ed.), *Handbook of reading research* (pp.353-394). New York: Longman.
- Barnett, M.A. (1988). Reading through context: How real and perceived strategy use affects L2 comprehension. *Modern Language Journal*, 72, 150-162.
- Bentler, P.M. (1985-2006). *EQS Version 6 for Windows* [Computer software]. Encino, CA: Multivariate Software.
- Bentler, P.M. (1995). *EQS structural equation program manual (Version 5.7)*. Encino, CA: Multivariate Software.
- Bentler, P.M. (2006). *EQS structural equation program manual (Version 6.0)*. Encino, CA: Multivariate Software.
- Bernhardt, E.B. (1986). Three approaches to reading comprehension in intermediate German. *Modern Language Journal*, 67, 111-115.
- Berhardt, E.B., & Kamil, M.L. (1995). Interpreting relationships between L1 and L2 reading: Consolidating the linguistic threshold and the linguistic independence hypothesis. *Applied Linguistics*, 16, 15-34.
- Block, E.L. (1992). See how they read: Comprehension monitoring of L1 and L2 readers. *TESOL Quarterly*, 26, 319-343.
- Brown, A.L., Bransford, R., Ferrara, R., & Campione, J. (1983) Learning, remembering, and understanding. In P. Mussen (Ed.), *Handbook of child psychology: Cognitive development* (4th Edition, Vol. 3; pp.77-166). New York: John Wiley and Sons.
- Byrne, B.M. (1994). *Structural equation modeling with EQS and EQS/Windows: Basic concepts, applications, and programming*. Thousand Oaks: Sage Publications.
- Carrell, P.L. (1984). Personality and individual writing processes. *College Composition and Communications*, 35, 285-300.
- Carrell, P. L. (1989). Metacognitive awareness and second language reading. *Modern Language Journal*, 73, 121-149.
- Carrell, P. L. (1991). Second language reading: Reading ability or language proficiency. *Applied Linguistics*, 12, 159-179.
- Carrell, P.L., Gajdusek, L., & Wise, T. (1998). Metacognition and EFL/ESL reading. *Instructional Science*, 26, 97-112.

- Chamot, A.U. (2005). Language learning strategy instruction: Current issues and research. *Annual review of Applied Linguistics*, 25, 112-130.
- Chamot, A.U., Kupper, L., & Interstate Research Associates, VA. (1989). Learning strategies in foreign language instruction. *Foreign Language Annals*, 22, 13-24.
- Cohen, A.D. (1998). *Strategies in learning and using a second language*. London and New York: Longman.
- Cohen, A.D. (in press). The coming of age of research on test-taking strategies. In J. Fox, M. Weshe, D. Bayliss, L. Cheng, C. Turner and C. Doe (Eds.), *Language testing reconsidered*. Ottawa: Ottawa University Press.
- Faerch, C., & Kasper, G. (1983). *Strategies in interlanguage communication*. London: Longman.
- Fuller, W.A., & Hidiroglou, M.A. (1978). Regression estimates after correcting for attenuation. *Journal of the American Statistical Association*, 73, 99-104.
- Gagné, E.D., Yekovich, C.W., & Yekovich, F.R. (1993). *The cognitive psychology of schooling learning*. New York, NY: HarperCollins College Publishers.
- Hacker, D.J. (1998). Self-regulated comprehension during normal reading. In D.J., Hacker, J. Dunlosky, & A.C. Graesser, *Metacognition in educational theory and practice* (pp.165-191). Mahwah, NJ: Lawrence Erlbaum Associates.
- Hymes, D. (1972). On communicative competence. In J.B. Pride, & J. Holmes (Eds.), *Sociolinguistics: Selected readings* (pp. 269-293). Harmondsworth, Middlesex: Penguin.
- Jöreskog, K.G., & Sörbom, D. (1988). *LISREL 7: A guide to the program and applications*. Chicago: SPSS.
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge: Cambridge University Press.
- Kline, R.B. (1998). *Principles and practice of structural equation modeling*. New York: Guilford.
- Maruyama, G.M. (1998). *Basics of structural equation modeling*. Thousand Oaks: Sage Publications.

-
- Mehrens, W.A., & Lehmann, I.J. (1984). *Measurement and evaluation in education and psychology (3rd Ed.)*. New York: Holt, Rinehart & Winston.
- McArdele, J.J., & Bell, R.Q. (2000). An introduction to latent growth models for developmental data analysis. In T.D. Little, K.U. Schnabel, & J. Baumert (Eds.), *Modeling longitudinal and multilevel data: Practical issues, applied approaches and specific examples* (pp.69-107). Mahwah, NJ: Lawrence Erlbaum Associates.
- McNamara, T. (1996). *Measuring second language performance*. London & New York: Longman.
- O'Malley, M.J., & Chamot, A.U. (1990). *Learning strategies in second language acquisition*. Cambridge: Cambridge University Press.
- Oxford, R. (1990). *Language learning strategies: What every teacher should know*. Boston: Heinle & Heinle Publishers.
- Oxford, R. (1996). Employing a questionnaire to assess the use of language learning strategies. *Applied Language Learning*, 7, 25-45.
- Pedhazur, E.J., & Schmelkin, L.P. (1992). *Measurement, design, and analysis: An integrated approach*. Hillsdale, NJ: Holt, Rinehart & Winston.
- Phakiti, A. (2003a). A closer look at gender differences in strategy use in L2 reading. *Language Learning*, 53, 649-702.
- Phakiti, A. (2003b). A closer look at the relationship of cognitive and metacognitive strategy use to EFL reading comprehension test performance. *Language Testing*, 20, 26-56.
- Pitts, M. M. (1983). Comprehension monitoring: Definition and practice. *Journal of Reading*, 26, 516-523.
- Pressley, M., & Afflerbach, P. (1995). *Verbal protocols of reading: The nature of constructively responsive reading*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Purpura, J.E. (1997). An analysis of the relationships between test takers' cognitive and metacognitive strategy use and second language test performance. *Language Learning*, 47, 289-325.

- Purpura, J.E. (1998). Investigating the effects of strategy use and second language test performance with high- and low-ability test takers: A structural equation modeling approach. *Language Testing*, 15, 333-379.
- Purpura, J.E. (1999). *Learner strategy use and performance on language tests: A structural equation modeling approach*. Cambridge: Cambridge University Press.
- Rumelhart, D. E. (1977). Toward an interactive model of reading. In S. Dornic & P. Rabbitt (Eds.), *Attention and performance* (Vol. 6, pp.573-603). New York: Academic Press.
- Salataci, R., & Akyel, A. (2002). Possible effects of strategy instruction on L1 and L2 reading. *Reading in a Foreign Language*, 14, 1-17.
- Schumacker, R.E., & Lomax, R.G. (1996). *A beginner's guide to structural equation modeling*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Segalowitz, N. (2003). Automaticity and second language development. In C. Doughty & M. Long (Eds.), *Handbook of second language acquisition* (pp. 382-408). Malden, MA: Blackwell.
- Song, X. (2004). *Language learning strategy use and language performance for Chinese learners of English*. Unpublished master's thesis. Ontario: Queen's University, Kingston, Canada.
- Song, X. (2005). Language learner strategy use and English proficiency on the Michigan English Language Assessment Battery. *Spain Fellow Working Papers in Second or Foreign Language Assessment*, 3, 1-26.
- Spielberger, C.D. (1972). Conceptual and methodological issues in anxiety research. In C.D. Spielberger (Ed.), *Anxiety: Current trends in theory and research*, Vol. 2 (pp.481-493). New York: Academic Press.
- Spolsky, B. (1995). *Measured words*. Oxford: Oxford University Press.
- Stevens, J. (1992). *Applied multivariate statistics for the social sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Stevens, J. (1996). *Applied multivariate statistics for the social sciences* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

-
- Tarone, E. (1998). Research on interlanguage variation: Implications for language testing. In L.F. Bachman, & A.D. Cohen (Eds.), *Interfaces between second language acquisition and language testing research* (pp.71-89). Cambridge: Cambridge University Press.
- Tseng W., Dörnyei, Z., & Schmitt, N. (2006). A new approach to assessing strategic learning: The case of self-regulation in vocabulary acquisition. *Applied Linguistics*, 27, 78-102.
- Wenden, A.L. (1991). *Learner strategies for learner autonomy*. Englewood Cliffs, NJ: Prentice-Hall.
- Wigglesworth, G. (2005). Current approaches to researching second language learner processes. *Annual Review of Applied Linguistics*, 25, 98-111.
- Yang, Y. (2002). Reassessing readers' comprehension monitoring. *Reading in a Foreign Language*, 14, 18-42.

**Appendix A The cognitive and metacognitive strategy
questionnaire**

Name-Surname: _____ Student ID: _____

Today's date: _____ Gender: [] male [] female Age: _____

No. of year learning English: ____ English Entrance Test Score: ____

Directions: A number of statements which people use to describe themselves when they were taking a reading test are given below. Read each statement and indicate how you thought during the test. Choose **1** (Never), **2** (Sometimes), **3** (Often), **4** (Usually), and **5** (Always).

	Your thinking	1	2	3	4	5
1.	I made short notes or underlined main ideas during the test.					
2.	I translated the reading texts and tasks into Thai.					
3.	I used pictures or titles of the texts to help comprehend reading tasks.					
4.	I used my own English structure knowledge to comprehend the text.					
5.	I spent more time on difficult questions.					
6.	I tried to understand the texts and questions regardless of my vocabulary knowledge.					
7.	I tried to find topics and main ideas by scanning and skimming.					
8.	I read the texts and questions several times to better understand them.					
9.	I used my prior knowledge to help understand the reading test.					
10.	I tried to identify easy and difficult test tasks.					
11.	When I started to complete the test, I planned how to complete it and followed the plan.					
12.	I was aware of what and how I was doing in the test.					
13.	I checked my own performance and progress while completing the test.					
14.	I attempted to identify main points of the					

	given reading texts and tasks.						
15.	I thought through the meaning of the test tasks/questions before answering them.						
16.	I was aware of which strategy to use and how and when to use it.						
17.	I corrected mistakes immediately when found.						
18.	I asked myself how the test questions and the given texts related to what I already knew.						
19.	I determined what the test tasks/questions required me to do.						
20.	I was aware of the need to plan a course of action.						
21.	I was aware of how much the test remained to be completed.						
22.	I tried to understand the questions adequately before attempting to find the answers.						
23.	I made sure I understood what had to be done and how to do it.						
24.	I was aware of my ongoing reading and test taking.						
25.	I kept track of my own progress to complete the questions on time.						
26.	I used multiple thinking strategies to help answer the test questions.						
27.	I made sure to clarify the goal and know how to complete it.						
28.	I checked my accuracy as I progressed through the test.						
29.	I selected relevant information to help me understand the reading texts and answer the test questions.						
30.	I carefully checked the answers before submitting the test.						