The Reading Matrix: An International Online Journal
Volume 19, Number 1, April 2019

Understanding L2 Reading Cognitive Processes: The Case of the L2 Reader's Goal

Abulfazl Mesgarshahr
Faculty of Foreign Languages and Literature, University of Tehran, Tehran, Iran

Seyyed Mohammad Alavi
Faculty of Foreign Languages and Literature, University of Tehran, Tehran, Iran

ABSTRACT

This study presents empirical evidence for the hypothesis that reading purpose may impact upon the cognitive processes that L2 readers use in comprehending a text. 25 advanced learners of English were put into two purpose groups – testing condition (TC) and learning condition (LC). The participants in the TC read two expository passages and answered 26 comprehension test items, while the LC participants read the same passages but without test items. Both groups were asked to verbalize their thoughts while performing the tasks. The data obtained from verbal reports, retrospective interviews, and the errors made during comprehending the passages or answering test items were collected and then analyzed. The results showed that readers pursuing different goals may engage in quantitatively as well as qualitatively different cognitive processes and strategies. The implications for L2 reading instruction and assessment are discussed.

INTRODUCTION

The expanse of research on first language (L1) reading is astonishingly vast. Multiple cognitive theories of reading have been proposed to explain what psychological or mental processes readers go through to comprehend a text (Graesser, Millis, & Zwaan, 1997; McNamara, & Magliano, 2009). In the domain of second/foreign language (L2) reading, however, a great deal of research is still needed to substantively understand the cognitive processes involved in text comprehension, thereby providing the foundation for developing L2 reading cognitive models (Grabe, 2009). Understanding L2 reading cognitive processes can significantly contribute to our knowledge about what the construct of L2 reading itself represents, what cognitive processing skills are involved in comprehending a text, what features can distinguish skilled from less skilled readers, how readers' strengths and weaknesses can be identified, and what instructional remedies can be suggested (Alderson, 2000; Grabe & Stoller, 2011).

In addition to understanding the substance of L2 reading cognitive processes, we also need to explore the different variables that might have the potential to manipulate these processes. A principled focus, through empirical research, on the reader, the text, and the interaction of the two might yield a number of variables that may impact upon the processing trajectories involved in L2 reading. This line of research can have important implications for both reading instruction and reading assessment. As regards the former, instructors and material developers will be informed of what cognitive processes or knowledge structures are common in various developmental stages of L2 acquisition, or what kind of cognitive processes are invoked by various textual features (e.g., rhetorical structures, vocabulary difficulty, grammatical complexity). In assessing reading, testers will better understand the psychology underlying test performance. This would allow them to have at their disposal a more comprehensive, theoretically supported picture of the construct they are to measure and make more valid interpretations of their assessment (Embretson, 2017; Messick, 1989; Nichols, 1994).

One variable pertaining to the reader that might influence the cognitive processes involved in L2 reading is the *reader's goal*. The reader's goal can be considered as an important component of metacognition, which is mostly responsible for monitoring and controlling cognitive processes and strategies (Hudson, 1996). Findings from research on L1 reading indicate that certain adjustments and modifications that align with the reading purpose are likely to be made to the cognitive processes and strategies used during reading (e.g., Linderholm & van den Broek, 2002; Narvaez, van den Broek, & Ruiz, 1999; van den Broek, Lorch, Linderholm, & Gustafson, 2001). That is, readers seeking different goals, namely reading in order to learn from the text or answer some text-related questions, might employ different cognitive processes to satisfy their particular goal. Reader's goal, although fairly well-researched in L1 reading, is a less explored variable in L2 reading research (Alderson, 2000; Grabe, 2009; Grabe & Stroller, 2011). The present study was set out to examine what impacts an L2 reader's goal might have on cognitive processes in text comprehension.

We will continue with a closer look at some L1/L2 reading research, with the main focus on the reader's goal and the concept of 'standard of coherence' (Linderholm & van den Broek, 2002) which is central to this study. This is followed by a brief discussion of the theoretical and practical issues surrounding 'verbal protocols,' a common technique in cognitive psychology for studying cognition (Sternberg & Sternberg, 2012). Next, we will describe in some detail the methodology used in this study for collecting verbal protocols. The article finally closes with a discussion of the results and the implications.

Reading Cognitive Processes

The current thought promoted by major cognitive models of L1 text comprehension would generally subscribe to the view that text understanding is brought about by the interaction of both bottom-up and top-down processes (McNamara, & Magliano, 2009; van den Broek et al., 2001; Verhoeven & Perfetti, 2008). Bottom-up or lower-level processes, such as decoding lexical items or understanding propositions at the sentence level, are typically automatic and consume only an infinitesimal amount of cognitive resources. These processes, which are largely realized through text-based inferences, constitute the critical phase of comprehension whereby the principal meaning connections between the propositions are established and the text-model of comprehension is constructed (Kintsch & Rawson, 2005; Perfetti, 2007; Zwaan & Singer, 2003). In contrast, top-down or higher-level processes, which expend more cognitive effort, operate to

integrate the amalgamated text information with relevant prior knowledge to allow the reader to assimilate the text, contributing to the formation of a coherent situation model of comprehension (Graesser, Singer & Trabasso, 1994; Zwaan & Brown, 1996). Research findings indicate that lower level processes are by no means less important than higher level processes and can be a critical factor in distinguishing less skilled from more skilled readers (Nassaji, 2003). Succinctly stated, text understanding consists of developing mental representations of the text by constructing a text model as well as a situation model of comprehension through generating a host of high- or low-level inferences.

The role of 'inference generation' is therefore pivotal to text understanding. Appropriately enough, a fair share of research on discourse understanding has been devoted to analyzing different types of inferences along with all the adjacent and distal reading components that might influence the nature of those inferences (e.g., Magliano, Trabasso, & Graesser, 1999; McNamara & Magliano, 2009). One such component, according to constructionist theories of discourse comprehension (Graesser & Kreuz, 1993), is the reader's goal. The idea is best captured by the Landscape model of reading developed by van den Broek and colleagues (van den Broek, Rapp, & Kendeou, 2005; Yeari & van den Broek, 2011). According to this model, readers maintain certain 'standards of coherence,' engage in varying inferential processes dictated by these standards, and eventually construct mental representations of the text that are consistent with the comprehension goal they pursue. For instance, if a reader's goal involves shallow processing, such as scanning through a text to find some piece of information, then the ensuing text comprehension, if any, will be spasmodic and superficial. Conversely, a higher standard of coherence will stimulate more inference generation and elaborative processing, and the resultant text representation will be more coherent with higher activation levels for key concepts (McNamara & Magliano, 2009). Hence, standards of coherence can be differentially sensitive to the readers' goal, and, consequently, have substantial effects on the reading process.

In research on L2 reading, however, there has been an intriguing scarcity of studies examining the degree to which reading goal can manipulate the reader's standards of coherence. The only few studies related to this area have been conducted mainly by Horiba (Horiba, 2000, 2013; Horiba & Fukaya, 2015). Horiba (2000), for instance, examined reader control with two types of tasks (reading freely and reading for more coherence) and found that the extent to which reader's goal can influence the reading cognitive processes might be subject to L2 proficiency. Readers with low L2 proficiency, she argued, might not have enough free cognitive resources, due to linguistic complexity of the text, to keep reading purpose in mind throughout reading.

In another study, Horiba (2013) compared three groups of Japanese EFL students who read an essay for expressions, read for image, and read for critique. By having participants think aloud to the essay, she found that resource allocations to linguistic processing, higher-level conceptual processing, and self-monitoring differed depending on the reading goal. Since the participants were of different L2 proficiency levels and recalled in either L1 or L2, findings also suggested that the reader's goal might have interacted with L2 proficiency and general comprehension skill. Moreover, Horiba and Fukaya (2015) investigated the impact of reading goal, topic-familiarity, and language proficiency on comprehension and learning. They used L2 proficiency and vocabulary acquisition measures along with recall protocols and found that the amount of content recall and incidental learning of new words varied depending on the particular goal the readers pursued. Although content recall and incidental vocabulary learning may not reflect the exact cognitive processes that the readers might have gone through, their results may

indicate that the final product of text processing is more likely to vary if readers hold different goals in mind.

To date, no study has been done to investigate the impact of reader's goal, while controlling for other intervening variables such as L2 proficiency, on the different cognitive processes identified by prior research. In order to fill this gap, the present study employed 'verbal protocol analysis' in the form of think-aloud methodology and retrospective interviews, along with analyzing readers' errors in text understanding, to explore in detail the cognitive processes that readers pursuing different goals use to comprehend a text. In the next section, we introduce the verbal protocol methodology.

Verbal Protocol

Verbal protocol, a common technique in cognitive psychology for studying complex cognitive processes, is believed to be well suited to the exploratory investigation of text understanding (van den Broek et al., 2001). A verbal protocol is an introspective method in which "the participants describe aloud all their thoughts and ideas during the performance of a given cognitive task" (Sternberg & Sternberg, 2012, p. 32). More precisely, the verbatim account of a problem solver thinking aloud while performing a cognitive task is called *concurrent verbal protocol*, also known as think-aloud protocol, whereas the problem solver's description of how the task was resolved, reported after accomplishing the activity itself, is labeled *retrospective verbal protocol* (Ericsson, 1988; Ericsson & Simon, 1993; Taylor & Dionne, 2000). Thorough reviews (Ericsson & Simon, 1993; Pressley & Afflerbach, 1995) of studies that used verbal protocols in their data collection procedure have concluded that verbal protocols elicited under specific conditions with regard to information-processing considerations can unveil the contents of the working memory during task performance and provide invaluable insight into cognitive processing.

The existing research literature on verbal protocol is not devoid of controversy, however (Alavi, 2005; Nisbett & Wilson, 1977; Smagorinsky, 1998). There are two major validity concerns regarding the use of verbal reports: reactivity and veridicality. Reactivity is concerned with the intrusiveness of thought verbalization with the undesirable result of disrupting the cognitive and thinking processes, while veridicality relates to whether verbal protocol accurately reflects the problem solver's cognitive processes (Ericsson, 2003). The argument is that concurrent reporting could introduce some superfluous cognitive load to the working memory (Just & Carpenter, 1992), thus interfering with the cognitive process it aims to describe. Similarly, the retrospective reports, despite not adding the cognitive weight that concurrent reports demand, might be subject to memory erosion and misconstrued by the researcher's goal and bias (Afflerbach & Johnston, 1984). In addition, it is argued that not all thought sequences can be reported since some cognitive processes are automatic and unconscious, thereby not amenable to verbalization (Ericsson & Simon, 1993).

With due attention paid to all of these limitations, experts in verbal protocol methodology have made a number of suggestions to improve the chances of obtaining non-reactive, veridical verbal reports. First, according to the Ericsson and Simon's (1993) information processing model, the respondents' explanations, not merely their thinking-aloud, might lead to reactive effects. Participants should therefore be discouraged from giving explanations that would otherwise be absent in the silent performing of the task (Chi, 1997; Ericsson, 2003; Ericsson & Simon, 1998). Second, since thinking aloud may not be a normal part of problem solving,

respondents should be given some standard instruction on what and how to verbalize. This instruction ought to be followed by some warm-up tasks similar to the ones in the experiment (Ericsson & Simon, 1998). Third, the tasks chosen for the experiment need to be of moderate difficulty (Afflerbach & Johnston, 1984; Leighton, 2004). Easy tasks would call upon automatic, unconscious processing, about which verbal reports are unable to disclose much, while exceedingly difficult tasks might thwart any cognitive effort to do the task itself, not to mention the thought reporting. Finally, conducting pilot studies is crucial (Ericsson & Simon, 1993). Given the cultural and individual idiosyncrasies of any group of subjects, a pilot study can draw a rough sketch of the procedural specificities and contextual subtleties, minimizing the likelihood of potential pitfalls for the tedious and time-consuming process of collecting and analyzing verbal protocols for the main study. Importantly, information can be obtained, through the pilot phase, on issues such as what kind of subjects, in terms of personality features, are more apt to produce adequate verbal reports, what tasks can be of moderate difficulty for the intended subjects, and what kind of instruction and probing can work best (Afflerbach & Johnston, 1984).

A further recommendation made by Taylor and Dionne (2000) encourages the complementary use of both concurrent and retrospective reports. One advantage, they believe, lies in juxtaposing the data obtained from both. Data compatibility of concurrent and retrospective accounts can illuminate or verify the interpretation of both, hence increasing validity, whereas their incongruity can pose challenging but insightful questions. Moreover, concurrent and retrospective reports have their own strengths and weaknesses, and, due to their very nature, might expose distinct cognitive processes and different forms of knowledge states relating to strategy use. For example, concurrent protocols can be more informative and less vulnerable to incorrect reporting (Ericsson & Simon, 1993), providing direct and unanalyzed reporting of thought as it occurs with minimal intervention of the researcher. Retrospective reports, on the one hand, liberate the respondent from the extra cognitive burden of thinking aloud during task performance, and, on the other hand, might be more appropriate for revealing information about meta-cognitive or executive knowledge by dint of the researcher's prompting (Taylor & Dionne, 2000). Therefore, triangulation of concurrent and retrospective reports can help improve the validity of the verbal protocols through both veridicality and comprehensiveness of the data.

THE PRESENT STUDY

The major purpose of this study was to examine whether an L2 reader's goal imposes any impact on the nature of cognitive processes involved in L2 reading. Specifically, this study aimed to examine the effects of two particular goals – reading to learn from a text versus reading to answer some passage-related questions such as in a reading comprehension test – on comprehension behaviors of L2 readers. This study, moreover, attempted to provide a detailed account of the cognitive processes underlying L2 text understanding. Text understating in a second language can be "a promising area to further test and develop theories and models of language comprehension" (Zwaan & Brown, 1996, p. 323).

A further point of significance for conducting such study comes from the importance attached to the situated nature of learning and understanding in accordance with recent developments in cognitive and educational psychology that call for broader ranges of student models and types of data (Mislevy, 1995). According to the 'cultural specificity of information

processing' premise of cognitive psychology (Rupp & Mislevy, 2007), the contents of learning provided, used, and transferred within a given group of people are peculiar to and characteristic of that group. Therefore, the educational decisions in a given context, related to either instructional or assessment issues, should be made based on the empirical research that take into account the immediate particularities of that context.

The main research question in this study is:

What effects might the L2 reader's goal have on the cognitive processes involved in reading comprehension?

METHOD

Participants

A total of 25 participants (15 females and 10 males) who were all Iranian learners of English as a foreign language took part in this study on a voluntary basis. All the participants were university students (freshmen and sophomores) from various majors (e.g., economics, engineering, education, translation). The first author, who was teaching English in advanced levels in a private language institute, informed his colleagues and his own students of the research project. Some students kindly consented to participate in the study, but, since people who are considered to be verbally superior might be able to talk their thought better (Chi, 1997; Ericson, 2003), we chose 25 of the volunteers who were seen as being more talkative. Because the two texts used in this study were taken from the reading module of IELTS (Appendix 1), all the participants had to be of advanced language proficiency level. They were all native speakers of Farsi (official language in Iran) who had not lived in any foreign country before.

Data Collection Procedure

The participants (N = 25) were randomly assigned into two reading purpose conditions: reading to learn from the text – learning condition (LC) – with 13 participants and reading to answer some passage-related questions in the form of a reading comprehension test – testing condition (TC) – with 12 participants. We opted for these two goals since these are the major purposes for which Iranian university students decide to learn English – to take international English language proficiency tests (i.e., TOEFL, IELTS, etc.) or to study and learn from texts written in English.

The verbal reports for each individual participant were collected by the first author in a two-hour session and tape-recorded to be later transcribed for further analysis. The researcher was present throughout the experiment while taking meticulous notes of the participant's verbal reports for more detailed investigation in the retrospective interviews. All the participants were first instructed on how to think aloud (Appendix 2) and then given a warm-up task. Depending on the purpose condition to which they belonged, the participants proceeded with the experiment differently. In order to simulate the real conditions of a reading comprehension test, the TC participants were given the two passages each with 13 comprehension test items all at once under the pre-specified time limitations (25 minutes for each passage and the questions while thinking aloud).

In contrast, the LC participants started with the same passages but without their comprehension questions. They were given fifteen minutes for each passage and asked to

imagine that they had to read the texts carefully in order to learn from them. Unlike TC participants, the LC participants were not told that the texts were taken from IELTS. For both group conditions, the thinking-aloud part was followed by the retrospective reports which included a semi-structured interview (Appendix 3) and an analysis of errors either in comprehension or in answering test items. The illustration in Figure 1 summarizes the data collection procedure.

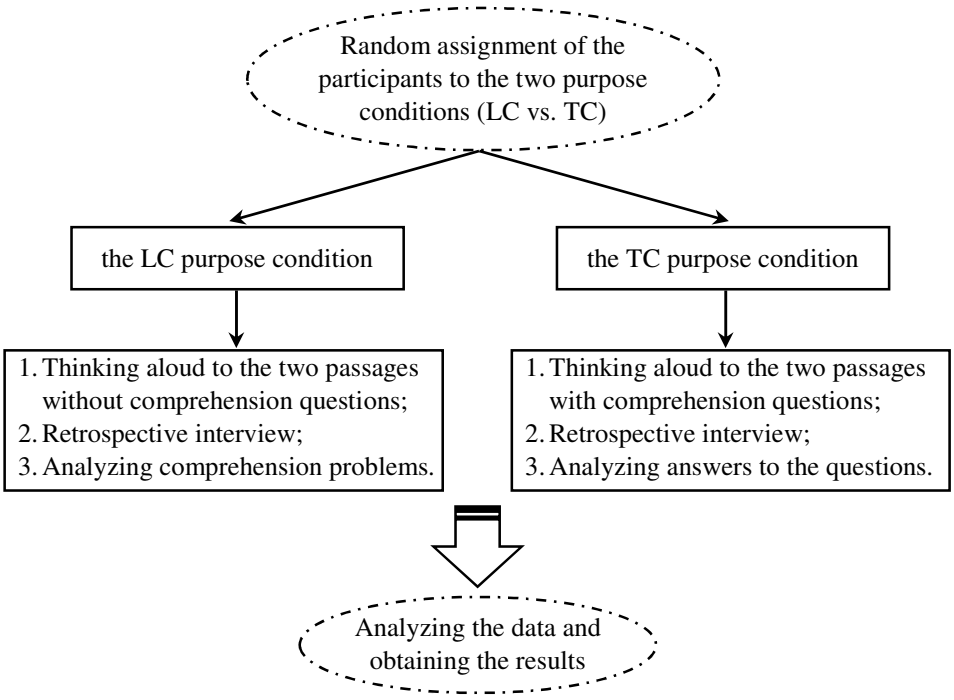

Figure 1. The data collection procedure

Materials

Different text types, mainly expository and narrative, are expected to trigger different cognitive processing mechanisms during comprehension (Graesser & McNamara, 2011; Horiba, 2000). Usually lengthy and informational, expository texts are decontextualized and primarily serve to inform the reader who typically does not have extensive background knowledge about the concepts and the ideas expounded on. Narrative types, in contrast, provide "a close correspondence to everyday experiences in contextually specific situations" (Graesser et al., 1994, p. 372) and permit more world knowledge to be at the reader's disposal, hence being less cognitively demanding and more accessible to processing (McNamara & Magliano, 2009). Since the intention was to explore the cognitive processes involved in L2 reading in situations reflective of university life, the materials used in this study were two sections from reading module of IELTS consisting of two expository texts each followed by 13 comprehension test items (Appendix 1).

The Pilot Study

A number of issues had to be resolved before proceeding with a large-scale collection of verbal reports. There was some concern over the tasks chosen (the reading module of IELTS),

such as whether they were of moderate difficulty for the intended participants as well as how many passages would be ideal for reading and verbal reporting. Also, since completing the tasks while verbalizing thought might take somewhat longer (Ericsson & Simon, 1993), the extra time needed for reading the passages and reporting the thought had to be determined. Moreover, we had to design a semi-structured interview composed of some revealing questions to elicit richer retrospective reports from respondents. We needed the observation of some participants doing the tasks to customize the questions we had in mind and come up with new ones. And finally, the procedure for thinking aloud had to be decided on. Different studies on comprehension have adopted various procedures for collecting verbal reports, such as having the subjects read a text sentence by sentence and verbalize their thought, or giving the subjects single cards with one sentence on each (e.g., Horiba, 1990; Linderholm & van Broek, 2002; Magliano & Millis, 2003; Zwaan & Brown, 1996), or even placing a colored dot at some preselected parts of the text (Crain-Thoreson, Lippman & McClendon-Magnuson, 1997). We had to ascertain, through a pilot phase, which method could serve the purpose of this study.

Seven participants took part in the pilot phase, upon conducting which we came to the following conclusions:

a) The participants almost unanimously agreed that the passages were neither too easy nor too difficult for them;
b) The typical reading module of IETLS consists of three sections, each with one passage followed by 13 or 14 test items. We observed that the participants felt exhausted after doing two sections. So we decided to conduct the main study with two sections of the reading module of IETLS to avoid the fatigue factor;
c) Twenty minutes is usually allocated for each section by the original test specifications. An extra time of five minutes for each section was considered adequate to provide enough time for the participants to do the tasks while thinking aloud;
d) Ten to twelve questions were developed to ask from the participants during the retrospective reporting;
e) Six out of seven participants did not feel comfortable verbalizing their thought after each sentence, seeing it as unnatural. Further, presenting the texts incrementally could distort the normal flow of reading, jeopardizing the ecological validity of the results. Therefore, it was decided, for the main study, to provide the participants with the whole text at once with verbalization left to the participants' discretion.

Moreover, it was observed that examining the errors the participants made either during text comprehension or while answering the questions could reveal more about their cognitive processing. So we decided to perform error analysis during the retrospective interviews.

Data Analysis

The think-aloud data along with the researcher's notes and the retrospective interviews were transcribed verbatim. Of varying types, the obtained data had to be analyzed based on two criteria: quantifiability and commonality. Most cognitive processes and strategies exposed by think-aloud protocols were quantifiable and common between the two purpose conditions. The analysis of these cognitive processes was performed according to the categories suggested by prior research (Cohen & Upton, 2006; Kern, 1994; Linderholm & van den Broek, 2002;

McNamara & Magliano, 2009; Weigle et al., 2013; Zwaan & Brown, 1996). These categories included:

- *backward inferences* (connecting current information to information that is previously encountered in the text);
- *forward inferences* (anticipating information that is yet to be described in the text);
- *associative inferences* (activating information from background knowledge that is associated with the current information in the text);
- *paraphrasing* (putting the current sentence into one's own words;
- *text repetition* (the verbatim reiteration of the current sentence or paragraph);
- *translation* (a mental reprocessing of L2 words, phrases, or sentences in LI forms while reading L2 texts);
- *deriving word meaning* (the attempt, failed or successful, at finding out the meaning of an unknown word);
- *evaluative comment* (expressing opinions about the text);
- *micro-metacognitive comments* (reflecting on one's own understanding or lack of understanding of the text at the word, phrase, or sentence level); and finally
- *macro-metacognitive comments* (reflecting on one's own understanding or lack of understanding of the text at paragraph or text level).

The coding of the protocols was done by the first author. Two other raters, who were Ph.D. candidates in TEFL, also performed the analysis using the same coding system for the protocols of five participants (one fifth of the whole data). The inter-rater agreement ratio was 87.70%. Any disagreement was resolved by discussion.

Additionally, a good portion of the data, gathered largely in the retrospective interviews, was not agreeable to quantification. Experiential familiarity with a certain test and background knowledge on the topic of a text, for instance, might prove advantageous to the test-taker and the reader, respectively. Also, as commonality concerns, some reading strategies or cognitive processes were specific to the TC and totally absent in the LC, as the latter had no test items to answer. The results of analyzing such data will be presented in the 'discussion' section below.

RESULTS

The independent variable in this study was reading purpose (LC vs. TC), and the dependent variables were cognitive processes and strategies (backward inferences, forward inferences, associative inferences, paraphrasing, test repetition, translation, deriving word meaning, evaluative comment, micro- and macro metacognitive comments). The data obtained were submitted to separate analyses of variance (ANOVAs). Since the number of participants was not equal in the two group conditions, type III sums of squares were used in all calculations to ensure the accuracy of *F* statistics. The results are shown in Table 1.

Table 1. Cognitive processes and strategies in reading for the two purpose conditions

Cognitive Processes and Strategies	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared

Backward Inference	336.748	1	336.748	20.782	.000	.475
Forward Inference	2.170	1	2.170	2.312	.142	.091
Associative Inference	1341.120	1	1341.120	34.509	.000	.600
Paraphrasing	34.348	1	34.348	3.177	.088	.121
Text Repetition	168.646	1	168.646	1.452	.241	.059
Translation	31.231	1	31.231	1.188	.287	.049
Deriving Word Meaning	474.954	1	474.954	29.077	.000	.558
Evaluative Comment	59.570	1	59.570	11.362	.003	.331
Micro Metacognitive Comment	4.813	1	4.813	.117	.735	.005
Macro Metacognitive Comment	392.351	1	392.351	43.677	.000	.655

As the results in Table 1 indicate, LC participants (M = 14.84, SD = 1.11) produced far more backward inferences than the TC participants (M = 7.50, SD = 1.16), $F(1,23) = 20.78$, $p < .05$. Associative inferences were also more common among LC participants (M = 20.07, SD = 1.72) than the TC participants (M = 5.41, SD = 1.80), $F(1,23) = 34.50$, $p < .05$. In deriving word meaning, the LC participants (M = 15.31, SD = 1.21) outperformed their TC counterparts (M = 6.58, SD = 1.16), $F(1,23) = 29.07$, $p < .05$. The LC participants (M = 6.92, SD = 0.63) also made more evaluative comments than their TC counterparts (M = 3.83, SD = 0.66), $F(1,23) = 11.36$, $p < .05$. And the final statistically significant difference between the LC (M = 12.84, SD = 0.83) and TC (M = 4.91, SD = 0.86) participants was found in making macro-metacognitive comments, $F(1,23) = 43.67$, $p < .05$. This quantitative analysis of the results shows that readers following different goals may go through distinct cognitive processes and strategies. A more detailed discussion of the findings is provided in the next section.

DISCUSSION

Standards of coherence: readers and test-takers

The main finding of this study was that L2 readers seeking different goals while reading a text are very likely to develop different standards of coherence, which, in turn, might affect their inference generation during reading. The results were also available for the differences between skilled and less skilled readers but these results will not be discussed here due to space limitations. Five out of the ten categories of reading cognitive processes examined in this study demonstrated statistically significant differences between the two purpose conditions:

1. *Backward inferences.* As discussed earlier, various types of inferential processes, constituting the text-based and knowledge-based models of comprehension, are amalgamated to accomplish text understanding. The text-based model of comprehension includes extracting word meanings to form idea units or propositions and then constructing intersentential relations to arrive at comprehension at the text level (Kintsch & Rawson, 2005). One inferential process that contributes to the formation of the text-based model is executed through generating backward inferences, whereby the information presented by the focal sentence is connected to the information encountered earlier in the text.

The results of this study showed that the TC participants produced fewer backward inferences than did the LC participants, suggesting that connecting ideas expressed in a text in order to achieve a globally coherent representation of the text may not be a major priority for the readers who aspire to answer some passage-related questions. The reason might be that the standard of comprehension that the TC participants develop is mostly driven by the questions following the texts. This may encourage shallow processing and superficial understanding accompanied by insufficient inference generation with a disjointed and less coherent textual representation. It was not uncommon for the TC participants to even skip some parts of the text. One TC participant asserted that:

> *I read the questions one by one and then scan the text to find the answer. Just as I find the answer for a question, I give up reading and go for the next question. If I feel that this part of the text doesn't contain my answer, why should I bother to read that?*

Contrastingly, the LC participants' high standards of comprehension sought proper relationships between the propositions explicitly provided in the text, hence stimulating more backward inferences. In this fashion, a globally coherent text-based model of comprehension is more likely to be achieved.

2. *Associative inferences.* Constructing a text-based model may not be sufficient for deep text comprehension to occur (Kintsch & Rawson, 2005). According to the constructionist theories of reading, deep comprehension requires active integration of textual information into related background knowledge and the goals of the reader, resulting in the formation of a situation model that reflects both local and global coherence (Graesser et al., 1994; McNamara & Kintsch, 1996). The information in the text is assimilated or elaborated on by prior knowledge that is not in the text via associative inferences (McNamara & Magliano, 2009), which are considered, by some researchers, to be the basis for the generation of other types of inferences (Kintsch, 1988). Seldom is a coherent situation model formed when the reader activates little knowledge beyond the information explicitly mentioned in the text, thereby presenting comprehension at a predominantly text-based level (McNamara & Magliano, 2009).

The LC participants in this study inclined to generate more associative inferences in order to mobilize more background knowledge and assimilate the textual information. This finding is also supported by the interesting observation that the LC participants tended to actively engage in and develop a friendly relationship with the text, attesting to which were a large number of examples whereby the LC participants exercised their knowledge-based judgments on the accuracy of the information delivered by the text. Upon reading the sentence 'This is a benign variation within circadian rhythms known as a chronotype,' one LC participant thought aloud, saying:

> *Hmm, night people are not sick?! Interesting! I thought people who are not kind of able to sleep at night but make up for it during the day are kind of sick. But here it says it's a benign variation. It's not dangerous! Cool!*

On the other hand, applying background knowledge and going beyond the text were by no means deemed to be necessary by the TC participants. The standards of coherence that the TC participants set may not have required them to integrate the newly encountered information in the text into the existing knowledge structures. Rather, these standards, driven by the reader's

goal to answer some passage-related questions, stipulated a somewhat simplistic representation of the text in a staccato rhythm. Correspondingly, higher level inferences, such as associative inferences, needed for achieving deep comprehension might not have been much sought-after. Moreover, TC participants, as test-takers, may intentionally prevent their background knowledge from having any impact on their understanding, as one participant said:

> *I always try to switch off everything I know about the topic of the text. Whatever the text says matters. I shouldn't choose any item as the right answer based on what I know.*

3. *Deriving word meaning.* Few would disagree that word identification is the sole ubiquitous and pervasive cognitive operation while reading (Perfetti, 2007). As long as being performed automatically, word identification can be considered as a lower-level, resource-cheap processing component that is indispensable to successful text model construction (Perfetti & Stafura, 2014). But multiple studies have argued that lexical access is more resource taxing in L2 than in L1 comprehension, especially when a less familiar word is encountered (e.g., Kern, 1994; Horiba, 2000; Zwaan & Brown, 1996). As for expository texts, there is a higher likelihood that readers confront unknown vocabulary. And the way the new words are tackled varies considerably depending on, inter alia, the reader's goal. However, Pressley and Afflerbach (1995) distinguished four phases through which the meaning of a new word might be treated or derived: (a) deciding whether the new word is worth expending any effort on; (b) carefully considering the co-text and the context of the new word; (c) arriving at a conclusion on a tentative meaning; and (d) subsequent evaluation of the derived word meaning.

The analysis of verbal reports indicated that the TC participants tried to derive the meaning of unknown words to a far less extent than did the LC participants, and that the reason for which the word meaning was speculated on differed. For instance, only when answering one item was predicated on knowing the meaning of a new word did the TC participants make any attempt to derive the meaning of the word, whereas LC participants chose to work on the meaning of a new word in order to enhance their text comprehension. One LC participant stated that:

> *I see, this word (circadian rhythm) pops up in the text repeatedly. So, if I want to understand the text, I should work out its meaning. Hmm, I have to find where the text talks about it for the first time. If I go back, um, yeah, it's here (reading the sentence). Yeah, right.*

Of importance, the LC participants' speculations on word meanings often turned out to be more accurate than those of TC participants, in that, the former, upon seeking more textual coherence, were more successful in performing subsequent evaluations of the derived word meaning, making meaning modifications highly probable. This is evidently reflective of the standards of coherence the participants had set.

4. *Evaluative comments.* Readers express their opinions about the text through evaluative comments (Linderholm & van den Broek, 2002), demonstrating their involvement and interest in the text. Evaluative comments (e.g., "This is exactly me!" "I don't buy that stuff!" "That's cool! Animals also can have this kind of feeling!") were made more frequently by the LC participants than by the TC participants. Upon reading the sentence 'The recommended course of action is to

follow an intense workout with a carbohydrate-rich breakfast,' one LC participant asserted that:

> *Oh, come on! You mean I've been doing this wrong for all these years?! No, no, these scientists say something different every day!*

This may indicate that the LC readers have reflected more on the textual information through the lens of their background knowledge in the direction of situation model construction to bring about deeper comprehension.

5. *Macro-metacognitive comments.* Through considering text understanding as a process operating at a local as well as a global level, this study maintained a distinction between micro- and macro-metacognitive comments, with the former referring to local comprehension and the latter to global comprehension. The verbal protocol analysis showed that the LC participants produced more macro-metacognitive comments than did the TC participants, whereas the two groups made roughly the same number of micro-metacognitive comments. One LC participant thought aloud:

> *So, until now, three kinds of brains. Reptilian, and limbic, and neocortex. This last one was for humans only. Yes! So, are they going to be other kinds? Let's continue.*

This is indicative of the LC participants' greater concerns over text comprehension at a sentence or text level, pointing again to the differences in the quality of the standards of coherence that the two purpose groups had developed.

These five reading cognitive processes (i.e., backward inferences, associative inferences, deriving word meaning, evaluative comment, and macro-metacognitive comments) showed statistically significant differences between the two purpose conditions. But treating the data exclusively from a quantitative perspective might run the risk of overlooking some important data. Several points can be made here. First, some cognitive processes, whose occurrences could be counted up, did not expose any statistically significant difference between the two purpose conditions, but they had qualitative potential. The difference in text repetition, for instance, between the two groups was statistically non-significant, while the nature of these repetitions and the reasons for which the participants decided to reread a part of the text were barely identical. Text repetitions for LC participants were mostly prompted by internal motivations, that is, the readers themselves chose which part of the text deserved a rereading in order to augment coherence and achieve deep comprehension, whereas the TC participants' repetitions were induced by some external drives – namely test items – that determined which parts had to be reread. Therefore, there might be some qualitative differences in the cognitive processes employed to apparently the same quantitative extent.

Second, not all cognitive processes were quantifiable. An important cognitive process which was common in the LC group but conspicuous by its absence in the TC group was 'hypothesis formation.' The LC participants usually had the tendency to go through the text by forming hypotheses, at the outset, about the main argument of the text. These hypotheses, often a mixture of both the text model and the situation model, were modified or even discarded and replaced as the readers proceeded with the text. This can be seen as further evidence that the LC participants attempted to meet their high standards of coherence.

Third, the presence of questions following the passages provoked certain cognitive processes or strategies. Since the LC participants had no questions to answer, these processes

were relevant only to the TC participants. One such process which we have called *realizing paraphrases* occurs when the test-taker recognizes that a meaning chunk in the text is repeated in another way in the test item. Also, the *process of elimination* is a test-taking strategy that involves "selecting an option even though it is not understood, out of a vague sense that the other options couldn't be correct" (Cohen & Upton, 2006, p. 37). *Locating information* can yet be another strategy related to the test items through which the test-taker attempts to pinpoint the likely location of the required piece of information. Successful application of this strategy depends upon choosing the appropriate keyword as well as the speed with which the test-taker is able to scan the text. These three cognitive processes or test-taking strategies may not create any bonds between the propositions provided in the text, hence no benefit for the text model. Nor do they link the text information to background knowledge to help form the situation model. Their only function concerns with finding the required information in the text to arrive at the right answer. This implies the superficiality of the standards of coherence developed by the TC participants.

In summary, the results of this study showed that the standards of coherence that the LC participants developed differed significantly from those set by the TC participants. The two groups of participants went through quantitatively as well as qualitatively different cognitive processes and strategies to meet those distinct standards. These results might have some certain implications for both reading instruction and reading assessment that are discussed in the next section.

CONCLUSIONS

The present study attempted to use verbal protocol methodology to examine what impacts reading purpose might have on the cognitive processes and strategies that L2 readers go through while comprehending a text. The results showed that readers seeking different goals were likely to develop different standards of coherence and engage in distinct cognitive processes accordingly. Among the ten categories that were chosen to represent L2 reading cognitive processes, five categories (backward inferences, associative inferences, deriving word meaning, evaluative comments, and macro-metacognitive comments) showed statistically significant differences between the two purpose conditions. These five categories also demonstrated some qualitative differences. For instance, the LC participants derived word meanings in order to enhance comprehension (motivated by the reader), while the TC participants derived word meanings to answer the related test item (dictated by the test).

The presence of questions following a text can have profound impacts on the cognitive processes in L2 reading. In actuality, the reading process takes its direction by the questions that follow the text. Some cognitive processes, such as realizing paraphrases, are employed only when readers have to answer some questions. Questions can further cause test anxiety and have debilitating effects on cognitive resources.

The results of this study might have important implications for both reading instruction and assessment. An immediate implication for reading instruction is that teachers and material developers should consider the main purposes for which a certain group of L2 learners in a given context want to acquire the reading ability. Major exams preparation courses and academic reading courses need to select and administer their tasks based on the impacts that readers' goal might have on the quality of cognitive processes employed during reading. Concerning L2

reading assessment, the results might imply that we need to distinguish between *readers* and *test-takers* when it comes to assessing L2 reading comprehension so as not to observe a *test-taker's comprehension behavior* in a testing situation and, based on the results, reach to erroneous conclusions about a *reader's competence* in reading. Therefore, given the importance that cognitive processes are increasingly assuming in the current views on construct validation (Embretson, 2017), it would be substantively unfounded to draw inferences on readers' competence based on their score in a reading comprehension test.

This study was concerned with the cognitive aspects of text comprehension. The analysis of verbal reports, however, revealed that the readers' affective state can also be an important variable in text understanding. One affective factor which was particularly prominent among the TC participants in this study was stress. The important point is that this affective factor might have some cognitive repercussions. In fact, stress and the ensuing anxiety have shown to have the potential to diminish the available working memory resources (Rai, Loschky, Harris, Peck & Cook, 2011; Sternberg & Sternberg, 2012). This is while 'working memory' in almost all models of reading comprehension is given a central role (e.g., Just & Carpenter, 1992; Kintsch, 1988; Yeari & van den Broek, 2011), and, as Hulstijn and Bossers (1992) have found, it assumes more importance for L2 reading at higher levels of language proficiency. The incapacitated working memory, due to affective pressures, along with the low standards of coherence that the TC participants had developed might account for their inferior status in inference generation compared to the LC participants in this study. Thus, further research is needed to explore how the reader's affective state can influence the cognitive processes involved in L2 text understanding.

There are some limitations that could restrict the generalizability of our findings. This study employed a group of Iranian learners of English with homogeneous cultural and L2 linguistic background. These participants thought aloud to two expository texts followed by a number of multiple-choice test items. Thus, more research is needed to explore cognitive processes in reading employed by readers with varying cultural and linguistic background. Moreover, in order to enhance our understanding of L2 reading cognitive processes, readers' comprehension of different text and test item types need to be examined by various methodologies, namely eye tracking or reaction time.

Abulfazl Mesgarshahr *is a Ph.D. candidate of TEFL at the university of Tehran. His main areas of research are cognitive aspects of reading comprehension and cognitive diagnostic assessment.*

Email: a.mesgarshahr@ut.ac.ir

Seyyed Mohammad Alavi *is full professor of TEFL at the university of Tehran. He has published a number of articles in national and international journals. His main areas of research include language testing and dynamic assessment.*

Email: smalavi@ut.ac.ir

REFERENCES

Afflerbach, P., & Johnston, P. (1984). Research methodology on the use of verbal reports in reading research. *Journal of Reading Behavior, 16*(4), 307-322.

Alavi, S. M. (2005). On the adequacy of verbal protocols in examining an underlying construct of a test. *Studies in Educational Evaluation, 31*, 1-26.

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

Chi, M. T. H. (1997). Quantifying qualitative analyses of verbal data: A practical guide. The *Journal of the Learning Sciences, 6*(3), 271-315.

Cohen, A. D., & Upton, T. A. (2006). *Strategies in responding to the new TOEFL reading Tasks.* (Monograph No. 33). Princeton, NJ: ETS.

Crain-Thoreson, C., Lippman, M. Z., & McClendon-Magnuson, D. (1997). Windows on comprehension: Reading comprehension processes as revealed by two-think aloud procedures. *Journal of Educational Psychology, 89*(4), 579-591.

Embretson, S. E. (2017). An integrative framework for construct validity. In A. A. Rupp & J. P. Leighton (Eds.), *The handbook of cognition and assessment: frameworks, methodologies, and applications* (pp. 102-123). West Sussex: Wiley-Blackwell Publishing.

Ericsson, K. A. (1988). Concurrent verbal reports on text comprehension: A review. *Text, 8*(4), 295-325.

Ericsson, K. A. (2003). Valid and non-reactive verbalization of thoughts during performance of tasks. *Journal of Consciousness Studies, 10*(9-10), 1-18.

Ericsson, K. A., & Simon, H. A. (1993). Protocol analysis: Verbal reports as data (Rev. ed.). Cambridge, MA: MIT Press.

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

Grabe, W., & Stoller, F. (2011). *Teaching and researching reading* (2nd ed.). Essex, England: Longman.

Graesser, A. C., & Kreuz, R. J. (1993). A theory of inference generation during text comprehension. *Discourse Processes, 16*, 145-160.

Graesser, A. C., & McNamara, D. S. (2011). Computational analyses of multilevel discourse comprehension. *Topics in Cognitive Science, 3*, 371–398.

Graesser, A. C., McNamara, D. S., Louwerse, M. M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, & Computers, 36*, 193-202.

Graesser, A. C., Millis, K. K., & Zwaan, R. A. (1997). Discourse comprehension. *Annual Review of Psychology, 48*, 163–189.

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review, 101*, 371-395.

Horiba, Y. (2000). Reader control in reading: Effects of language competence, text type, and task. *Discourse Processes, 29*(3), 223-267.

Horiba, Y. (2013). Task-induced strategic processing in L2 text comprehension. *Reading in a Foreign Language, 25*, 98–125.

Horiba, Y., & Fukaya, K. (2015). Reading and learning from L2 text: Effects of reading goal, topic familiarity, and language proficiency. *Reading in a Foreign Language, 27*, 22–46.

Hudson, T. (1996). Assessing second language academic reading from a communicative competence perspective: Relevance for TOEFL 2000 (TOEFL Monograph Series Report No. 4). Princeton, NJ: Educational Testing Service.

Hulstijn, J. H., & Bossers, B. (1992). Individual differences in L2 proficiency as a function of L1 proficiency. *European Journal of Cognitive Psychology, 4*(4), 341–353.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension. *Psychological Review, 99*, 122–149.

Kern, R. G. (1994). The role of mental translation in second language reading. *Studies in Second Language Acquisition, 16*, 441-461.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review, 95*, 163-182.

Kintsch, W., & Rawson, K. A. (2005). Comprehension. In Margaret J. Snowling & Ch. Hulme (Eds.), *The Science of Reading: A Handbook*, (pp. 209-226). Malden, MA: Blackwell Publishing.

Leighton, J. (2004). Avoiding misconception, misuse, and missed opportunities: the collection of verbal reports in educational achievement testing. *Educational Measurement: Issues and Practice, 23*(4), 6-15.

Linderholm, T., & van Broek, P. (2002). The effects of reading purpose and working memory capacity on the processing of expository text. *Journal of Educational Psychology, 94*(4) 778-784.

Magliano, J. P., & Millis, K. K. (2003). Assessing reading skill with a think-aloud procedure and latent semantic analysis. *Cognition and Instruction, 21*(3), 251-283.

Magliano, J. P., Trabasso, T., & Graesser, A. C. (1999). Strategic processing during comprehension. *Journal of Educational Psychology, 9*, 615–629.

McNamara, D. S., & Magliano, J. (2009). Toward a comprehensive model of comprehension. In B. Ross (Ed.), *The psychology of learning and motivation* (Vol. 51, pp. 297–384). New York, NY: Elsevier Science.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 1–103). New York, NY: American Council on Education/Macmillan.

Narvaez, D., van den Broek, P., & Ruiz, A. B. (1999). The influence of reading purpose on inference generation and comprehension in reading. *Journal of Educational Psychology, 91*, 488-496.

Nassaji, H. (2003). Higher-level and lower-level text processing skills in advanced ESL reading comprehension. *The Modern Language Journal, 87*(2), 261-276.

Nichols, P. D. (1994). A framework for developing cognitively diagnostic assessments. *Review of Educational Research, 64*(4), 575–603.

Nisbett, R. E. & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84*, 231-259.

Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading, 11*(4), 357-383.

Perfetti, C., & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading, 18*, 22-37.

Pressley,M., & Afflerbach, P. (1995). *Verbal protocols of reading: The nature of constructively responsive reading.* Mahwah, NJ: Erlbaum.

Rai, M. K., Loschky, L. C., Harris, R. J., Peck, N. R., & Cook, L. G. (2011). Effects of stress and working memory capacity on foreign language readers' inferential processing during comprehension. *Language Learning, 61*(1), 187-218.

Smagorinsky, P. (1998). Thinking and speech and protocol analysis. *Mind, Culture, and Activity, 5*(3), 157-177.

Sternberg, R. J., & Sternberg, K. (2012). *Cognitive psychology* (6th ed.). Wadsworth: Cengage Learning.

Taylor, K. L., & Dionne, J. P. (2000). Accessing problem-solving strategy knowledge: The complementary use of concurrent verbal protocols and retrospective debriefing. *Journal of Educationa1 Psychology, 92*, 413-435.

van den Broek, P., Lorch, R. F., Jr., Linderholm, T., & Gustafson, M. (2001). The effects of readers' goals on the generation of inferences. *Memory & Cognition, 29*, 1081–1087.

van den Broek, P., Rapp, D. N., & Kendeou, P. (2005). Integrating memory-based and constructionist approaches in accounts of reading comprehension. *Discourse Processes, 39*, 299–316.

Verhoeven, L. & Perfetti, C. (2008). Advances in text comprehension: Model, process, and development. *Applied Cognitive Psychology, 22*, 293–301.

Yeari, M., & van den Broek, P. (2011). A cognitive account of discourse understanding and discourse interpretation: The landscape model of reading. *Discourse Studies, 13*(5) 635-643.

Zwaan, R. A., & Brown, C. M. (1996). The influence of language proficiency and comprehension skill on situation-model construction. *Discourse Processes, 21*, 289-327.

Zwaan, R. A., & Singer, M. (2003). Text comprehension. In A. C. Graesser, M. A. Gernsbacher, & S. R. Goldman (Eds.), *Handbook of Discourse Processes* (pp. 83-121). Mahwah, New Jersey: Routledge.

APPENDIX 1

Some general information about the two passages used for thinking aloud

Passage 1

Title:	"Making Time For Science"
Word count:	763
Flesch Reading Ease score:	49.442 (fairly difficult to read)
Flesch-Kincaid Grade Level:	10.575
Test item format:	seven "True," "False," or "Not Given" items; and six multiple choice items.
Test item types:	Explicitly/implicitly stated information; main idea.

Passage 2

Title:	"The Triune Brain"
Word count:	834
Flesch Reading Ease score:	47.178 (difficult to read)
Flesch-Kincaid Grade Level:	10.741
Test item format:	9 matching items; and 4 summary completion items.
Test item types:	Explicitly/implicitly stated information; main idea.; summary completion

Notes:

- The two passages along with their test items were retrieved from https://takeielts.britishcouncil.org/prepare-test/practice-tests/reading-practice-test-1-academic/reading-passages.
- The two passages were analyzed by Coh-Metrix (Graesser, McNamara, Louwerse & Cai, 2004) provided by www.cohmetric.com.

APPENDIX 2

The instructions given to the participants

The instructions for the LC group:

> Please imagine you are sitting in your room alone and want to study two texts that are important for you to understand and learn the content. While you are reading the text, you should bring to your tongue whatever the text brings to your mind. Whatever thought you may think is trivial can be significant for us. If you keep silent for some time, I will say: "Keep talking."

The instructions for the TC group:

> Please imagine you are going to take a very important exam. There are two texts each with 13 test items, taken from IELTS. While you are reading the text and answering the questions, you should bring to your tongue whatever comes to your mind. Whatever thought you may think is trivial can be significant for us. If you keep silent for some time, I will say: "Keep talking."

APPENDIX 3

The questions used in semi-structured interviews

Questions used in retrospective interview for the TC group

1. Did you make any time plans for doing the task (reading the text and answering the items)?
2. What mainly accounts for the difficulty of the texts/test items? (vocabulary, grammar, length, time, etc.)
3. Which strategy best helps you overcome the difficulties in comprehension (rereading, translation, relating the focal sentence to the whole text, etc.)?
4. Which strategy best helps you answer the test items successfully (looking for key words, going through the test items before reading the text, reading the text paragraph by paragraph, etc.)?
5. Do you ever try to summarize one paragraph to see what it talks about and then go to another one?
6. Did you read the whole text? Did you skip any part of the text? Why?
7. Doesn't moving back and forth between the text and the questions disrupt your comprehension of the text?
8. How different would your reading be if there weren't any questions to answer?
9. Do you read the text mainly to answer the items or to comprehend the text?
10. Does thinking aloud have any effect on your reading comprehension?

Questions used in retrospective interview for the LC group

1. Did you make any time plans for doing the task (reading the text and answering the items)?
2. Did you employ any strategies to understand the text better (i.e., skimming the whole text first to get the main idea and then reading more carefully)?
3. What mainly accounts for the difficulty of the texts? (vocabulary, grammar, length, etc.)
4. Which strategy best helps you overcome the difficulties in comprehension (rereading, translation, relating the focal sentence to the whole text, etc.)?
5. Do you ever try to summarize one paragraph to see what it talks about and then go to another one?
6. Did you read the whole text? Did you skip any part of the text? Why?
7. How different would your reading be if there were some questions to answer just like a reading comprehension test?
8. Does thinking aloud have any effect on your comprehension?