International Journal of Language and Linguistics
2022; 10(1): 5-15
http://www.sciencepublishinggroup.com/j/ijll
doi: 10.11648/j.ijll.20221001.12
ISSN: 2330-0205 (Print); ISSN: 2330-0221 (Online)

The Influence of Computer-Based Reading Assessment on Chinese EFL Learners' Use of Reading Strategies and Their Reading Performance

Yangyang Yu

School of Foreign Languages, Shanghai Sanda University, Shanghai, China

Email address:
tomatoeggs0209@163.com

To cite this article:
Yangyang Yu. The Influence of Computer-Based Reading Assessment on Chinese EFL Learners' Use of Reading Strategies and Their Reading Performance. *International Journal of Language and Linguistics*. Vol. 10, No. 1, 2022, pp. 5-15. doi: 10.11648/j.ijll.20221001.12

Received: December 14, 2021; **Accepted**: December 25, 2021; **Published**: January 12, 2022

Abstract: Computer-based assessment has been increasingly adopted in second language testing with the development of computer technology. Previous studies have been devoted to comparing computer-based assessment with paper-based assessment in terms of test results, but little is known about how the test process may be changed by the new format. With a focus on the assessment of reading, this study aimed to investigate the influence of computer-based format on the use of strategies during the reading process together with the final reading performance. It adopted case study as the methodology, recruiting twenty Chinese EFL learners at the university level as the participants. Two reading tasks were taken from College English Test Band 6 (CET6), respectively held in a paper-based format and a computer-based format. Questionnaires on use of reading strategies were adapted from previous studies to measure students' strategy use during reading. Selected participants were interviewed after the reading tasks and the questionnaire to supplement quantitative data. It was found that students tended to perform better in a computer-based setting than a paper-based one, and that reading strategies can influence the final performance and may be used differently in the two settings. The possible reasons and pedagogical implications were then discussed and explored.

Keywords: Computer-Based Assessment, Reading Strategies, Reading Performance

1. Introduction

Information technology has revolutionized the area of second language teaching and testing in the past two decades. As electronic resources are prevailing in almost every aspect of life, from work and study to entertainment and recreation, digital literacy, namely the ability to access information in an online environment, has become an increasingly critical concern for language education [3]. In response to this concern, language testing has started a conversion in the testing medium, chiefly from the conventional paper-based format to the newly-developed computer-based format. Some high-stake English exams like TOEFL, GRE and GMAT have been employing computer in test delivery, administration and scoring for years; others like Business English Certificate (BEC) and College English Test (CET) in China have started to adopt computer as an alternative for paper and pencil as their testing medium. It seems that computer-based language assessment has become the trend in a technology era, receiving much attention from experts and scholars.

Previous research has been conducted on the comparability between paper-based testing (PBT) and computer-based testing (CBT) in assessing language, delving into the affordances and limitations of CBT from the perspective of test results [11]. However, few studies have been devoted to how computer, as a test medium that establishes a digital context, may specifically influence learners' cognitive process and demonstration of a certain language skill during a test. Given reading on the screen entails a different way of information processing from that in reading on a printed page [7], it can be speculated that language learners would change their reading habits under a CBT condition, which may influence the final reading performance. Use of reading strategies can be one of such reading habits that is worth examining, because it has proved an important factor in

reading achievements for EFL learners with various first languages. Therefore, exploring how the CBT setting may affect reading strategies and performance is likely to be a good starting point to complement previous research, understand washback effects of CBT, revise test design and enhance language teaching in a world of digital reading.

The present study, accordingly, intends to investigate the possible effect of computer-based reading assessment on use of reading strategy and final reading performance, particularly in the context of Mainland China. There are two basic reasons for selecting the target context. On one hand, China is now experimenting computer-based language test for the establishment of a standardized English language test; on the other hand, despite large numbers of EFL learners in Mainland China, the area of English testing is relatively under-researched, unable to provide sufficient empirical evidences for the reform of the testing system and the improvement of language teaching. Exploring the effect of computer-based reading assessment on both test results and test process may offer justification for the test reform and give inspiration to language teaching in China. Due to limited resources, it is barely possible for this study to involve a large sample or a sophisticated experiment design. Hence, a case study is adopted here with both quantitative and qualitative approaches, trying to reveal part of the process underlying CBT, and generating implications in research and pedagogy. Considering the general background discussed above, the study attempts to answer the following three research questions:

1. *Is there any significant difference between a computer-based reading task (CBRT) and a paper-based reading task (PBRT) in reading performance and strategy using for Chinese EFL learners?*
2. *What reading strategies significantly correlate with reading performance in a CBRT/PBRT?*
3. *How do Chinese EFL learners perceive strategy using in a CBRT/PBRT?*

2. Review of Literature

Before presenting the case study and unveiling the results, it is necessary to first review literature in relevant aspects so as to clarify the specific background for this study. Previous research findings in computer-based language assessment, reading performance and strategy using, and English testing in Mainland China are cited and discussed to develop the research niche of this study, laying foundation for the ensuing experiment design and data interpretation.

2.1. Computer-Based Language Assessment

As CBT is usually juxtaposed with PBT, a frequently employed approach to examining the former is comparing it with the latter in light of validity, reliability, efficiency, and other key qualities of testing [11]. Comparability studies have been conducted to contrast the two test formats in assessing four language skills, with a major focus on whether CBT brings tests of higher quality. It is generally agreed that CBT largely enhances testing speed and efficiency, but opinions on its validity vary among researchers [6].

It is pointed out that CBT measures a different construct from PBT in assessing language skills, introducing computer familiarity as a new variable that affects test outcomes [22]. To this extent, construct validity, the quality that a test monitors what it intends to monitor, is compromised. Since CBT requires test takers to have a certain level of computer familiarity for task items, it is not only the target language skills but also the computer ability that is tested and reflected in the final score. A test taker may be inferior to another in the result of a language test, not because s/he has a lower level of a certain language skill but because s/he is lacking in computer ability. This language test, therefore, fails to generate fair and trustworthy results, losing construct validity.

A counterargument to CBT's limitation in construct validity is that computer familiarity bears a close relationship with language ability when listening to, talking to, reading on, and writing on computer is prevalent in the real world [14]. The authenticity of computer-based language task seems to make computer familiarity a somewhat legitimate part of test construct. Take the writing skill for an example. If a language learner is expected to type on computer with a target language, his/her familiarity with computer should be also be monitored in the language test, as s/he cannot perform a real-world task without this ability. Similar justification can also be found in assessing listening [4], speaking [2], and reading [9], in that the computer-based format is somewhat inseparable from the authentic use of language. Nevertheless, it should also be noted that reading generally requires lower computer familiarity than the other language skills like writing, because it normally involves only clicking with a mouse instead of typing on a keyboard. With a focus on reading, the present study may avoid much interference from technical ability.

Despite deliberate evaluation of CBT compared with PBT, past studies center more on test result than on test process. Little is known about test takers' cognitive process during a test. According to Sawaki [28], even with similar results across the two different testing media, CBT construct is not the same as PBT construct, for the presentation mode of input in testing alters the cognitive process, but not necessarily engenders salient difference in the outcomes. Although the cognitive process during a test is not as measurable as final performance, instruments like eye movement devices, post-test interviews, questionnaires may help to uncover the invisible process [18]. The present study utilizes the latter two as the main source of data that describe the cognitive activity respectively in CBT and PBT settings.

2.2. Reading Performance and Strategy Using

Reading is a fundamental receptive skill that generates "comprehensible input" for the practice and development of the other three language skills [16]. There are two reading types: reading text aloud and reading comprehension. The one normally tested in second language exams is the latter, which refers to understanding the meaning of a given text with no need to speak out words or phrase [17]. The present study also

focuses on reading comprehension, not only because it comprises a regular section in language testing, but also because it tends to be integrated into testing of other language skills, indicating its critical role in second language acquisition.

Koda points out that reading comprehension is a complex cognitive activity that follows the steps of "decoding linguistic information, integrating the extracted information, and synthesizing text information with prior knowledge" [15]. Based on this process of reading, researchers have summarized key factors influencing L2 reading performance, which include L2 linguistic knowledge, background knowledge and use of strategy [21]. It is found that better knowledge of L2 vocabulary and grammar, and higher familiarity with text topic are usually conducive to L2 reading achievements [34, 1, 27]. However, the relationship between reading performance and strategy using is slightly complicated. Reading strategies are methods employed by readers to construct meaning effectively. Generally, appropriate strategy using facilitates reading comprehension and task performance; while specifically, not all strategies are helpful to a certain reading task, and some of them even bring negative impact [10]. In this sense, identifying useful reading strategies may enhance learners' strategic competence and therefore reading performance.

Reading strategies can be classified based on different standards. A commonly adopted way of classification is proposed by O'Malley and Chamot's, who divided strategies into two main categories: metacognitive strategies and cognitive strategies [23]. The former category includes specific strategies like planning, monitoring, self-evaluation; and the latter includes translating, analyzing, inferring, summarizing, predicting, note taking and others. With reference to such classification, the relationship between reading performance and strategy using has been examined in the traditional PBT setting. Utilizing questionnaires and interviews as research instruments, Phakiti suggests although both cognitive and metacognitive strategies have a positive influence on reading performance, successful test takers tend to employ more metacognitive strategies [26]. Sun revised the experiment design of Phakiti and took into consideration the proficiency level of readers, confirming the previous findings and indicating specific reading strategies relevant in a reading test for readers of both high and low language proficiency [30]. These studies form a basis for further research on reading performance and strategy using in the CBT setting, making it reasonable to hypothesize that a similar relationship may be existing and researchable.

On the other hand, a computer screen creates a different reading environment from that of a printed page, requiring readers to scroll rather than turn over, or mark with a mouse rather than with a pencil [7]. The physical differences between electronic and printed texts, therefore, result in different ways of processing information, which possibly trigger changes in reading performance and strategy using. Smuck investigated students' reading achievements across PBT and CBT settings, finding that those with high digital literacy generally performed better in CBT than in PBT [29]. This result pinpoints that the presentation mode of CBT brings positive impact in a digital context. However, to what extent such variance may be explained by different strategy using remains unknown. Although it is admitted that most reading strategies are transferrable from the paper-based to the computer-based format [24], the specific usage of them respectively in the two formats has not been clarified. Moreover, Park and Kim identified strategies unique to online contexts, and some of them are applicable to the testing condition [25]. In this sense, it is possible that use of reading strategies varies with testing media, and that reading performance also change in a corresponding way. Such possibility has barely been researched but may generate inspiring thoughts in language testing and teaching.

2.3. English Testing in Mainland China

As a compulsory subject from primary to tertiary education, English is highly valued in Mainland China where numerous learners study it for academic and professional purposes. Various English tests are developed to gauge the language proficiency of test takers, and provide certificate for their language ability. While China somehow lags behind in researching the effect of information technology on language education, Chinese test designers follow the trend of a technology era in practice, and increasingly combine the use of computer into English tests in forms of computer-based assessment, computer adaptive testing and automated scoring [19].

The initial trial of CBT in China is made in College English Test (CET), a high-stake exam taken by thousands of college students each year. There are two test levels, CET4 (lower) and CET6 (higher), which both test four language skills. Tertiary institutions used to set passing CET6 as a graduation requirement for students. This rule has now been abolished, but CET remains an important indicator of language proficiency, and almost an inevitable proof for job application and further education in Mainland China [35]. CET was originally available only in paper-based version, but the computer-based counterpart was developed in 2008 with growing awareness of digital literacy. Jin notes that while computer familiarity and test anxiety in the CBT version probably interfere with the construct validity of CET, increased computer skills may alleviate such interference, and even facilitate test performance [12]. However, due to insufficient empirical evidences supporting this argument, the CBT version of CET has been neither widely implemented nor accepted by most test takers. More research efforts are needed to explore the underlying features of CET in the CBT setting for the sake of test implementation and revision. Besides, considering China intends to establish a new standardized English test in a computer-based form [13], research on computer-based CET can be beneficial to the preparation of this new test.

In summary, past studies suggest a research gap to be addressed. First, comparisons between PBT and CBT have been made far more in test results than test process. Second,

although use of strategies has proved a crucial factor for reading comprehension performance in PBT, it remains unknown whether the relation exists in CBT. Computer-based format turns out to have an influence on final reading performance, but its effect on use of strategy during reading still needs exploration. Third, English testing in mainland China has encountered difficulty in implementing CBT, largely due to insufficient empirical evidences of the validity in CBT, and lack of justifiable ways to improve this new test format. Considering these issues, the present study compares PBT and CBT not only in final reading results, but also in use of strategies during the reading process in the context of China. It may reveal specific influence of computer as a medium on reading assessment, complementing previous research outcomes, and providing implications for language testing in China.

3. Methodology

The present study employed case study as its methodology. With limited resources, it is hardly possible for this study to obtain a large sample population. Nevertheless, the case study method allows "holistic and in-depth" exploration about research questions with a small sample size examined in a detailed way [33]. Therefore, the study followed the construct of case study, closely analyzing the behaviors and ideas of a small number of participants to reveal the possible situation of a larger learner groups. The three research questions mainly concern reading performance and use of strategies as variables to be measured. Although performance and strategy use may be measured quantitatively through tests and questionnaires, participants' specific perception about strategy using under different testing media cannot be fully illustrated with only quantitative data, which may impede exploring the effect of computer-based assessment on the reading process. So the study adopted mixed-methods, collecting and analyzing both quantitative and qualitative data in the framework of case study. The quantitative data came from reading tasks and questionnaires, directly showing reading performance and frequency of strategy using; the qualitative data could be retrieved from interviews on the reading process after the tasks, indicating participants' detailed feelings towards different test modes and preference in use of strategies. The analysis of the two types of data might exhaust the participant cases, uncover a relatively complete picture of the general situation, and generate a deeper understanding of the research questions.

3.1. Participants

The participants were 20 university students studying English as a foreign language in Shanghai, Mainland China. Before the experiment, it had been confirmed that the participants had already achieved the proficiency level required by CET6. In other words, they would not fail most CET6 questions because of great deficiency in grammar or vocabulary. This prerequisite was to reduce the influence of linguistic knowledge as a confounding factor for reading performance. To complete the two reading tasks and the following questionnaires on reading strategies, all the participants should have no physical difficulty in reading on a screen or a printed page, and possess at least very basic knowledge about reading strategies.

3.2. Instruments

3.2.1. Reading Tasks

The computer-based reading task (CBRT) and the paper-based reading task (PBRT) were chosen from the reading comprehension section of CET6 (2016), which is a nationwide English test in mainland China, assessing the English proficiency of undergraduates and postgraduates. There were two reasons for using CET6 reading task: first, CET6 is a carefully designed public exam with high validity and reliability, which is supposed to reflect readers' true reading performance; second, CET6 still has not completed the conversion from a paper-based test to a computer-based one, and the results of the present study may be relevant to such practice.

Each task consisted of two passages followed by five multiple-choice questions that tested readers' comprehension of the text. The full score for each task was 10. The study adopted a within-subject design, which meant all the participants went through the two tasks. This helped to lessen the impact of individual differences which would cause variance in study results. The reading materials and questions were selected from equivalent parts of two different test papers for CET6 on the same date, and hence were of similar difficulty.

The PBRT was presented in a traditional paper version, while the CBRT was presented with an online platform. Although the interface of this CBRT was not perfect due to technical restriction in this study, it could perform basic function of presentation and allowing selection.

3.2.2. Questionnaire on Use of Reading Strategies

The questionnaire in the present study was adapted from Phakiti (2003) and Sun (2011), both of which investigated the relationship between reading performance and strategy using with a questionnaire consisting of statements describing use of different strategies during reading. The participants were required to rate the statements according to a five-point Likert scale: 1 for never, 2 for rarely, 3 for sometimes, 4 for usually, 5 for always, so as to show their frequency of strategy using. The present study made some adaptation of the original questionnaires, basically by deleting the unnecessary repeated items, adding relevant items, and revising expression. Besides, the strategies concerned in this questionnaire were transferrable across paper-based and computer-based formats for the sake of comparison.

The two tasks each were followed by a questionnaire of the same contents, but in different presentation media. In other words, there was a paper-based questionnaire, and a computer-based questionnaire with a different order of statements to ensure reliability. Table 1 shows categories of reading strategies examined in this study, and the numbers of their corresponding statements respectively in the two

questionnaires. See Appendix for the complete paper-based questionnaire. The computer-based questionnaire could be accessed together with CBRT through a link (https://www.wjx.cn/jq/21460181.aspx). Demographic data of the participants were collected at the beginning of the paper-based questionnaire. The participants were required to give their name again in the computer-based task to match their paper-based performance.

Table 1. *Reading strategy categories and their corresponding item numbers in the questionnaires.*

Main categories	Specific categories	Number in PBRT	Number in CBRT
Metacognitive strategies	Advanced organization	1, 2, 3	1, 11, 15
	Selective attention	5, 7, 15	3, 5, 13
	Directed attention	6	14
	Self-evaluation of performance	17, 29	12, 19
	Self-evaluation of problem	30	18
	Comprehension monitoring	22	24
	Task monitoring	20, 21, 28	9, 10, 20
	Self-management	19, 26	21, 25
Cognitive strategies	Skimming	4	2
	Prediction	8	27
	Analyzing	9, 27	16, 17
	Inferring	10, 25	4, 22
	Translating	11	28
	Summarizing	12	30
	Repetition	14	7
	Guessing	16, 18	6, 26
	Elaboration	13, 23	8, 29
	Note taking	24	23

3.2.3. Interview Questions

Six participants were interviewed about their feelings after the two reading tasks. They were selected based on their average score of the tasks, 2 from higher achieving, 2 from average achieving, 2 from lower achieving. The interview was flexible and conducted within 3 minutes online, involving questions like:

1. How do you feel about the reading task presented on the screen?
2. Have you been taught how to use strategies to cope with a reading task?
3. What strategies do you prefer to employ in PBRT/CBRT?

3.3. Procedures

The Participants were asked to go through the same procedures. First, they took a PBRT adapted from a CET6 reading section, and then completed a questionnaire about strategies used in the previous reading task. On the next day, they took a CBRT of similar difficulty, and completed the same questionnaire on computer, but with a different order of statements. The two tasks and questionnaires were held on different days to avoid participants simply copying answers in questionnaires of the same contents. For each task, there was a suggested time limit, though not strictly practiced. Finally, 6 selected participants were interviewed on their feelings about the two tasks. The focus was on their use of reading strategies during tasks.

3.4. Analysis of Data

Quantitative data were collected from reading questions and strategy questionnaires, then recorded and processed by SPSS. Paired sample t-tests were used to verify whether the scores of two tasks were statistically different and whether the results of questionnaires were also different. Pearson correlation test was run to confirm the relationship between reading performance and strategy using, and identify specific strategies significantly correlate with reading performance in PBRT/CBRT. Qualitative data were obtained from transcribed interviews. Content analysis was used to process these data to reveal participants' specific feelings about their use of strategy in the two reading tasks.

4. Results

The result section contains three parts. The first part is about the reading performance in the PBRT and the CBRT, which is mainly revealed by task scores. The second part centers on the use of reading strategies under the two task conditions, presenting results from the questionnaires. The third part analyzes the interview contents pertinent to the research focus in this study.

4.1. Reading Performance

A total of 20 participants took the two reading tasks. Table 2 shows the descriptive data of the task scores. The average score in the PBRT (6.65/10) is lower than that in the CBRT (7.80/10), showing the participants achieve better performance in the CBRT on average. However, the comparison of means alone cannot generalize this findings to other cases. Thus, a paired-sample t-test was performed and presented in Table 3. The difference is significant at the 0.05 level, which means if the sig. (2-tailed) is lower than 0.05, there can be a significant difference in reading performance between the CBRT and the PBRT; while if the sig. (2-tailed) is higher than 0.05, the difference can be insignificant. In this sense, the reading performance in the PBRT is significantly different from that in the CBRT (Sig.=0.000<0.05). In other

words, there is at least 95% possibility that the CBRT would generate higher average score than the PBRT would do.

Table 2. *Descriptive data of two task scores.*

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	ScorePBRT	6.6500	20	1.98083	.44293
	ScoreCBRT	7.8000	20	1.43637	.32118

Table 3. *Paired-sample t-test of two task scores.*

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	ScorePBRT - ScoreCBRT	-1.15000	.93330	.20869	-1.58680	-.71320	-5.510	19	.000

Note: The difference is significant at the level of 0.05.

4.2. Use of Reading Strategies

All the participants rated statements describing use of reading strategies in a questionnaire after each reading task. The five-point Likert scale is employed to indicate the frequency of using certain strategy: 1 for never, 2 for rarely, 3 for sometimes, 4 for usually, 5 for always. Table 4 shows the findings gathered from the two questionnaires, including means of strategy using frequency in the PBRT and the CBRT, and paired-sample t-test of frequency differences between the two tasks. Although the using frequency of every specific category of strategy differs in means between the PBRT and the CBRT, the t-test results suggest significant difference exists only in 4 specific categories: selective attention (sig.=0.007<0.05), task monitoring (sig.=0.001<0.05), skimming (sig.=0.022<0.05), note taking (sig.=0.000<0.05). Considering the means in relation to the t-test results, selective attention (m1=3.30>m2=2.55) and note taking (m1=3.10>m2=1.40) are significantly more frequent in the PBRT; while task monitoring (m1=2.65<m2=3.45) and skimming (m1=2.20<m2=3.00) are significantly more frequent in the CBRT.

Table 4. *Frequency means and paired-sample t-test results of strategy using in the two tasks.*

Main categories	Specific categories	Frequency mean in PBRT (m1)	Frequency mean in CBRT (m2)	T	Sig. (2-tailed)
Metacognitive strategies	Advanced organization	3.60	3.55	.203	.841
	Selective attention	3.30	2.55	3.000	.007
	Directed attention	2.65	2.90	-.925	.367
	Self-evaluation of performance	3.55	3.20	.645	.527
	Self-evaluation of problem	2.80	2.55	1.314	.204
	Comprehension monitoring	1.75	2.05	-1.674	.110
	Task monitoring	2.65	3.45	-3.760	.001
	Self-management	3.40	3.20	1.703	.297
Cognitive strategies	Skimming	2.20	3.00	-2.491	.022
	Prediction	1.90	2.00	-.525	.606
	Analyzing	2.95	3.30	-1.161	.260
	Inferring	3.60	3.65	-.295	.772
	Translating	2.50	2.75	-1.000	.330
	Summarizing	2.15	1.75	1.799	.088
	Repetition	2.65	2.40	1.314	.204
	Guessing	3.35	3.75	-1.506	.148
	Elaboration	2.65	2.75	-.490	.629
	Note taking	3.10	1.40	8.233	.000

Note: The difference is significant at the level of 0.05.

Pearson correlation was performed to test the relationship between reading performance and use of specific strategies. The value of Pearson correlation (r) is between -1 and 1. The correlation between two variables can be positive (0<r<1), zero (r=0), and negative (-1<r<0). Positive correlation exists where two variables change in the same direction, negative correlation exists where two variables change in the opposite direction, and zero correlation shows there is no relation between two variables at all. The correlation is significant at either the 0.05 level or the 0.01 level, which means only when the sig. value is lower than 0.05 or 0.01 can a correlation be regarded as significant, generalizable to other cases. In this study, it turns out strategies that positively correlate with reading performance under both conditions include selective attention (r1=0.749, sig.1=0.000; r2=0.594, sig.2=0.006), self-evaluation of performance (r1=0.609, sig.1=0.004; r2=0.734, sig.2=0.000), task monitoring (r1=0.664, sig.1=0.001; r2=0.743, sig.2=0.000), self-management (r1=0.792, sig.1=0.000; r2=0.793, sig.2=0.000), skimming (r1=0.696, sig.1=0.001; r2=0.657, sig.2=0.002), analyzing (r1=0.596, sig.1=0.006; r2=0.661, sig.2=0.002), inferring (r1=0.631, sig.1=0.003; r2=0.704, sig.2=0.000), and guessing

(r1=0.548, sig1.=0.012; r2=0.474; sig.2=0.035). Directed attention (r1=- 0.665, sig.1=0.01; r2=- 0.465, sig.2=0.039) and translating (r1=- 0.705, sig.1=0.001; r2=- 0.787, sig.2=0.000) tend to have a negative correlation with reading performance in both CBT and PBT. The use of the other strategies does not appear to vary significantly with reading performance.

Table 5. *Pearson correlation between reading performance and use of strategies.*

Main categories	Specific categories	Pearson Correlation for PBRT (r1)	Sig. 1 (2-tailed)	Pearson Correlation for CBRT (r2)	Sig. 2 (2-tailed)
Metacognitive strategies	Advanced organization	.383	.095	.301	.184
	Selective attention	.794**	.000	.594**	.006
	Directed attention	-.665**	.001	-.465*	.039
	Self-evaluation of performance	.609**	.004	.734**	.000
	Self-evaluation of problem	.401	.079	.438	.054
	Comprehension monitoring	.279	.234	.348	.133
	Task monitoring	.664**	.001	.743**	.000
	Self-management	.792**	.000	.793**	.000
Cognitive strategies	Skimming	.696**	.001	.657**	.002
	Prediction	.290	.215	.369	.110
	Analyzing	.596**	.006	.661**	.002
	Inferring	.631**	.003	0.704**	.000
	Translating	-.705**	.001	-.787**	.000
	Summarizing	.369	.110	.414	.070
	Repetition	.321	.168	.429	.509
	Guessing	.548*	.012	.474*	.035
	Elaboration	.386	.093	.379	.099
	Note taking	.079	.742	.044	.855

* The correlation is significant at the 0.05 level.
**The correlation is significant at the 0.01 level.

4.3. Interview

Six participants were selected to be interviewed after the two reading tasks. Two came from the higher achieving group, named as P1 and P2, two from the average achieving group, named as P3 and P4, and two from the lower achieving group, named as P5 and P6. Table 5 shows the task results of these selected interviewees.

Table 6. *Selected interviewees' scores.*

Interviewee	PBRT Score	CBRT Score	Average score
P1	10	10	10
P2	9	10	9.5
P3	8	7	7.5
P4	7	7	7
P5	4	4	4
P6	2	3	2.5

When asked about the feelings toward the reading task presented on the screen, four participants (P2, P3, P4, P6) reported that they were quite familiar with the digital presentation, and the two rest (P1, P5) said that they found it acceptable to read on the computer, though preferring the paper-based format. P2 mentioned his experience of preparing for the computer-based TOEFL, stating that he was used to taking an online language test. P3, P4 and P6 all attributed their comfort with digital texts to much online reading either required by their study or done for entertainment. On the other hand, P1 put that she was more comfortable with a paper-based task, because reading on screen was somehow a burden to her eyes; and P5 said he did not take an online test before, and still needed some time to adjust himself to this test mode.

As for use of reading strategy, all the interviewees reported to have received instructions in reading strategies from language teachers as well as exercise books. Most of the strategies listed in the questionnaires were said to be familiar to them. Nevertheless, the actual use of reading strategies varied with participants. P1, P2 and P4 noted that they were trained in using strategies to solve reading comprehension questions, and thus employed many of the listed strategies consciously in the two reading tasks. P3 and P5 realized they used quite a few strategies after finishing the questionnaires. For them, strategy using was more unconscious than conscious. P6 admitted that despite having knowledge of reading strategies, she did not know how to use them, and preferred to directly solve the reading questions on her pace.

All the interviewees except P6 were then asked about whether they used reading strategies differently in the PBRT and the CBRT. P1 and P5 found it difficult to mark reading passages and take notes in the computer settings, which they considered relevant in their reading process. P5 especially claimed that he could hardly attend to the contents without underlining key words with a pencil. It can be indicated that they might decrease use of strategies involving marking and writing in the CBRT. By comparison, others did not consider being unable to mark a problem in the computer setting. P3 admitted using software to mark digital texts in non-test settings, but she thought marking may not be a necessity for her. P2 pointed out that locating information in the text and checking answers to the questions became easier when scrolling up and down on a webpage. P4 found her strategy using quite the same in the two tasks.

5. Discussion

This section tries to interpret the above results in response

to the research questions raised at the beginning of the study. Findings from previous research are reviewed and compared to what has been found in this study. Pedagogical implications are finally generated.

5.1. Differences Between a CBRT and a PBRT in Reading Performance and Strategy Using for Chinese EFL Learners

The analysis of task scores confirms there is a significant different between a CBRT and a PBRT in reading performance for Chinese EFL learners. It is indicated that learners tend to perform better with computer as the testing medium of a reading task. This finding supports previous research that shows the digital context created by computer-based assessment may have a positive influence on L2 learners [8, 29]. However, it seems to contrast with some comparability studies that suggest computer-based language tests generate results not significantly different from those in paper-based ones [20], or that better performance comes with the paper-based setting [5]. The inconsistency may result from different language skills examined in the studies and various experiment contexts. For example, the construct of a writing test deviates from that of a reading test, and therefore the computerized tests may not have the same influence on readers. Besides, backgrounds of L2 learners may also account for the variance of results, for various groups of learners have their respective traits in computer familiarity, language needs and learning habits. Thus, the generalizability of a study comparing the two test modes may be largely limited because of possible factors that affect the study result. As far as the present study is concerned, it may be concluded that EFL learners at the university level in China are likely to achieve higher scores in a CBRT than a PBRT. According to the interviews, the reason probably lies in much exposure to computer technology and online reading in study and daily life, which makes reading on the screen an acceptable, and even "enjoyable and comfortable" experience for test takers [9]. Given Krashen's affective filter hypothesis, such positive emotions may facilitate language input being processed by individuals. From this perspective, the conversion to computer-based language test may be feasible at least for testing reading among university learners in China. Testing of other abilities should be confirmed in further studies.

The comparison between results from the two questionnaires shows that the use of some reading strategies in a CBRT is also significantly different from that in a PBRT. It is found that the using frequency of task monitoring, selective attention, skimming and note taking are likely to vary with the testing medium, and that the rest of strategies may not be that influenced. A PBRT may lead to more use of selective attention and note taking; while a CBRT may come with more use of task monitoring and skimming. The differences can be traced back to the characteristics of two presentation modes. Computer allows much convenience in quickly locating information through scrolling, while paper and pencil enables marking that may serve as reminder during reading. Whether the different use of these reading strategies can explain the gap in performance between the two task conditions is still contingent on the correlation between specific strategies and reading performance, which will be discussed later in the second research question. Besides, it should be noted that the strategies examined here are limited to those applicable to the testing context. In fact, more possibilities in reading strategies can be brought by computer in other contexts. Park and Kim concludes access to hypermedia, use of computer skills and devices, scrolling up and down, moving back and forth as reading strategies peculiar to computer-based reading [25]. The influence of these unique strategies in contexts other than language testing may be explored in future research.

5.2. Reading Strategies Significantly Correlating with Reading Performance in a CBRT/PBRT

The Pearson Correlation test suggests there are metacognitive and cognitive strategies significantly correlating with reading performance under both computer and paper settings. Specifically, selective attention, self-evaluation of performance, task monitoring, self-management, analyzing, inferring, guessing are strategies with a positive correlation with reading performance, directed attention and translating are strategies negatively correlating with reading performance, and the rest of the listed strategies do not show significant correlation. In other words, more use of the first seven strategies may be conducive to reading comprehension, while more use of directed attention and translating may impede understanding an English text. Such findings lend support to Unzueta and Barbetta's argument that use of reading strategies generally benefits a reading task, no matter the task is computer-based or paper-based [31]. The specific strategies showing a significant correlation also echo with those pointed out by Sun, which adopted a similar questionnaire to investigate Chinese English majors' use of reading strategies for paper-based TEM 4 (Test for English Majors Band 4) [30]. A minor difference lies in that Sun also found advanced organization and summarizing with certain correlation with reading performance, probably because a TEM 4 task may require higher level of intensive reading skills. Therefore, the features of a reading task itself can be a factor influencing the relation between certain strategies and final reading achievements.

Among the reading strategies proving to closely correlate with reading performance in both a CBRT and a PBRT, task monitoring, skimming and selective attention turn out to be used differently between the two settings. The three strategies all positively influence the reading performance, but task monitoring and skimming are more frequently used in a CBRT, and selective attention occurs more in a PBRT, as is previously discussed in the first research question. Such findings jointly point to the speculation raised at the beginning of the study that the change of testing medium may alter the final reading performance partly through influencing the use of reading strategies. To put it more clearly, two chief facts from the study support this speculation. First, more use of task monitoring and skimming coexists with higher task score in a CBRT. Second, the rest of strategies, except selective attention, are either used similarly in the two task settings or bears no significant correlation with reading performance, which

means they have little influence on the differences of performance between a CBRT and a PBRT. It should be noted that selective attention may particularly benefit reading performance in a PBRT, but the effect may be inferior to the positive influence brought by task monitoring and skimming to a CBRT. There are many reasons for better performance in a CBRT, and this study just indicates the possibility that different use of certain reading strategies in the digital context can be one of them. Phakiti regards strategy using as a mediator between one's internal linguistic knowledge and external language contexts, suggesting the function of strategies in helping language learners improve their performance based on their own language ability and outside linguistic input [26]. It is therefore possible that the computerized input may engender better reading performance after being mediated by use of certain strategies (e.g. task monitoring and skimming) in a CBRT. The more use of selective attention in a PBRT may also perform such function, albeit not directly told from the final score in this study.

5.3. Chinese EFL Learners' Perception of Strategy Use in a CBRT/PBRT

The interview contents reveal the selected participants' specific perceptions about the two reading tasks, especially use of strategies during the reading process. First, it can be suggested that previous exposure to computer-based texts largely influence interviewees' feelings toward a CBRT. Much exposure to the digital context usually leads to comfort with computer-based assessment. Nonetheless, individual situations, for example one's physical condition, may change his or her perceptions about a CBRT. Secondly, considering the interviews in relation to the achievement levels of the interviewees, it can be indicated that in both settings, higher achieving learners tend to employ more reading strategies intentionally, while the lower achieving ones may employ fewer. Most interviewees integrated strategies into the reading tasks either intentionally or unconsciously, which to some extent, benefits their reading performance. In contrast, lack of strategy using may account partly for the relatively unsatisfactory task results of P6. This finding is consistent with the statistical analysis of task scores and strategy using frequency. Moreover, although strategy using seems to be transferrable from a PBRT to a CBRT, the minor different use of strategies can be perceived mostly in skimming, monitoring, marking and taking notes, echoing with the four strategies by the questionnaires. The interviews reveal that learners may feel encouraged to skim and monitor task in the computer-based setting, and thus achieve better understanding of a reading passage and higher completion of a comprehension task. Marking for selective attention and taking notes encouraged in the paper-based setting are also considered beneficial to reading, but the positive effect is less reported in the interview. It can be inferred that learners may be susceptible to the impact of different strategy using induced by paper or computer. Nonetheless, how the impact is passed on to the final reading performance also depends on individual situations. After all, some learners may be restricted in marking key words to give selective attention while still remaining high achieving in a CBRT. This study just points to a general tendency that computer may encourage strategies like skimming and task monitoring, which are likely to facilitate reading comprehension.

5.4. Pedagogical Implications

This study generates pedagogical implications mainly from three aspects. First, it is necessary to explicitly teach reading strategies in second language class, especially those that prove to positively affect reading performance in both paper-based and computer-based settings. According to Worrell et al., increased use of strategy usually brings more accurate reading comprehension regardless of the presentation mode [32]. Therefore, explicit instructions in strategy using can build up learners' strategic competence which will ultimately benefit their reading performance. Secondly, computer-based reading assessment can be feasible for second language testing in China. Previous comparability studies endeavor to justify or argue against computer as a testing medium by showing that results of a computer-based test and a paper-based test are similar or different. However, it is worth pondering whether integrating computer into language test is just about finding an equivalent to paper and pencil. Jin regards the computer-based language assessment as a response to the contextual concerns in today's China, where digital literacy is highly valued and computer familiarity is less an obstacle [12]. The fact that learners may achieve better reading performance with more use of certain strategies supports that computer caters for needs and wants from test takers and should be provided as a choice in testing medium. Besides, computer-based assessment may be modified to allow marking electronic texts, increasing the use of selective attention, which is now less employed in CBT but positively correlates with final performance. What's more, given language testing and teaching are closely related, computer technology should also be combined into language class to prepare learners for computer-based tasks in language as well as other fields. Computer-assisted language learning has proved effective for learners, but remains to be developed in mainland China.

6. Conclusion

The present study investigated the effect of computer-based reading assessment on Chinese EFL learners' reading performance and use of strategies. It was found that learners tended to perform better in a computer-based setting than a paper-based one. Learners' increased exposure to electronic resources and growing digital literacy may explain the favored performance in computer-based reading assessment. A few reading strategies were identified, which usually benefit final performance in both settings, but it turned out change of the testing medium may alter the using frequency in some of them, including skimming and task monitoring, which appeared more in the computer-based task, and selective attention, which occurred more in the paper-based task. There is possibility that more use of strategies like skimming and task monitoring contributes to better performance in the computer-based setting. With these findings, this study may complement previous research which focuses more on test results than on test process when comparing the two media, and generate useful pedagogical implications for language testing and teaching

in the Chinese context.

However, there are several limitations in the present study. First, the number of participants is relatively small, failing to include more diverse participants, and thus restricting the generalizability of the findings. Second, the interface of the computer-based task is rather simple, which can hardly simulate a formal language test. Third, the question type of the reading tasks is limited to multiple choice questions, which may not fully reveal the comprehension of a passage. Fourth, there is only indirect evidence supporting the speculation that computer as a testing medium may change the reading performance through influencing the test takers use of strategies. More direct evidence needs to be found to verify the speculation. Therefore, further research may improve the design flaws in the present study by including more participants, upgrading task contents and interface, and designing methods that directly monitor the reading process and generate more objective data about strategy using to supplement self-reported information. Of course, the changes brought by the computer-based format to the process of testing other language skills may also be explored.

Appendix – Questionnaire on Reading Strategies

Directions: The following statements can be used to describe your use of strategy during the previous reading task. Please read each statement and choose the scale number best describing how you thought during the task.

Please note the numbers indicate frequency: 1 (Never), 2 (Sometimes), 3 (Often), 4 (Usually) and 5 (Always).

Table 7. *Questionnaire on use of reading strategies.*

Name	**Age**	**Gender**	**Year of study**				
	Your thinking		**Your rating**				
1	I set plans on how to complete the reading task		1	2	3	4	5
2	I was aware of the objective of the reading task		1	2	3	4	5
3	I knew what needed to be done and how to do it		1	2	3	4	5
4	I skimmed the text quickly to have a general understanding of it		1	2	3	4	5
5	I paid attention to the questions and memorized them before reading the given text		1	2	3	4	5
6	I knew what to read closely and what to ignore to gain the main idea of the text		1	2	3	4	5
7	I read the text quickly to find out the relevant information for the questions		1	2	3	4	5
8	I predicted the content of the upcoming part of text while reading		1	2	3	4	5
9	I examined the relationship between the given text and the questions		1	2	3	4	5
10	I attempted to understand the implicit meaning of the given text		1	2	3	4	5
11	I translated what I had read into Chinese		1	2	3	4	5
12	I summarized the key points of the text		1	2	3	4	5
13	I connected what I read with my prior experience		1	2	3	4	5
14	When I failed to understand certain text or questions, I reread them to increase understanding		1	2	3	4	5
15	I tried to mark key words and sentences while reading		1	2	3	4	5
16	I guessed the meaning of unknown words or phrases based on the context or text clues		1	2	3	4	5
17	I carefully checked the answers before submitting the task		1	2	3	4	5
18	I guessed the meaning of unknown words based on their roots or affixes		1	2	3	4	5
19	I adjusted my reading speed based on different reading purposes		1	2	3	4	5
20	I was aware of how much of the task remained to be completed		1	2	3	4	5
21	I kept track of my own progress to complete the questions on time		1	2	3	4	5
22	I checked whether I understood the contents I had read before reading on		1	2	3	4	5
23	I related my background knowledge to text information to better understand the given text		1	2	3	4	5
24	I took notes when reading to help memorize useful information		1	2	3	4	5
25	I spotted transitional words (eg. first, second, however, but, because and etc.) to help understand the inner logic of the text.		1	2	3	4	5
26	I distinguished between easy and difficult questions and spent more time on difficult ones		1	2	3	4	5
27	I tried to break down long sentences to figure out their meanings		1	2	3	4	5
28	I corrected mistakes immediately when I thought I misunderstood the text or questions		1	2	3	4	5
29	I evaluated whether the reading plans were achieved		1	2	3	4	5
30	I tried to find out my weakness in the reading task, and thought about how to improve my reading efficiency		1	2	3	4	5

References

[1] Akbari, Z. (2014). The Role of Grammar in Second Language Reading Comprehension: Iranian ESP Context. *Procedia - Social and Behavioral Sciences, 98*, 122-126.

[2] Amengual-Pizarro, M., & García-Laborda, J. (2017). Analysing test-takers' views on a computer-based speaking test. *Profile: Issues in Teachers' Professional Development, 19* (Suppl. 1), 23-38.

[3] Bulger, M. E. (2006). Beyond search: A preliminary skill set for online literacy. *Transliteracies Project*. Retrieved from http://transliteracies.english.ucsb.edu.

[4] Coniam, D. (2006). Evaluating Computer-Based and Paper-Based Versions of an English-Language Listening Test. *ReCALL, 18* (2), 193-211.

[5] Cumming, A., Kantor, K. Baba, U. & James, M. (2006). *Analysis of discourse features and verification of scoring levels for independent and integrated prototype written tasks for next generation TOEFL*. Princeton, NJ: Educational Testing Service.

[6] DOOEY, P. (2008). Language testing and technology: problems of transition to a new era. *ReCALL, 20* (1), 21-34.

[7] Douglas, D. (2010). *Understanding Language Testing.* B. Comrie & G. Corbett (Eds.). New York: Routledge.

[8] Eno, L. (2011). *Comparing the Reading Performance of High-achieving Adolescents: Computer-based Testing versus Paper/pencil*, ProQuest Dissertations and Theses.

[9] Hosseini, M., Abidin, M. & Baghdarnia, M. (2014). Comparability of Test Results of Computer Based Tests (CBT) and Paper and Pencil Tests (PPT) among English Language Learners in Iran. *Procedia - Social and Behavioral Sciences 98*, 659–667.

[10] Huang, J. & Nisbet, D. L. (2015). Reading strategy use and reading proficiency of EFL students in China. *The Asian Journal of Applied Linguistics, 2 (3)*, 202-212.

[11] Jamieson, J. (2005). Trends in computer-based second language assessment. *Annual Review of Applied Linguistics, 25*, 228-242.

[12] Jin, Y. (2011). Fundamental concerns in high-stakes language testing: The case of the College English Test. *Journal of Pan-Pacific Association of Applied Linguistics, 15* (2), 71-83.

[13] Jin, Y. (2017). Construct and content in context: implications for language learning, teaching and assessment in China. *Language Testing in Asia, 7* (12), 1-18.

[14] Jin Y. & Yan, M. (2017). Computer literacy and the construct validity of a highstakes computer-based writing assessment. *Language Assessment Quarterly, 14* (2), 101-119.

[15] Koda, K. (2005). *Insights into Second Language Reading: A crosslinguistic Approach.* Cambridge University Press.

[16] Krashen, S. (1982). *Principles and practice in second language acquisition.* Oxford: Pergamon.

[17] Lee, H. (2013). Thinking Matters: Inferencing in ESL Reading Lessons. *TESOL Journal, 4* (4), 717-742.

[18] Lim, H. (2014). *Exploring the Validity Evidence of the TOEFL IBT Reading Test from a Cognitive Perspective.* Retrieved from ProQuest Dissertations and Theses.

[19] Liu, H., You, X., Wang, W., Ding, S., & Chang, H. (2013). The Development of Computerized Adaptive Testing with Cognitive Diagnosis for an English Achievement Test in China. *Journal of Classification, 30* (2), 152-172.

[20] Mason, B. J., Patry, M. & Bernstein, D. J. (2001). An examination of the equivalence between non-adaptive computer-based and traditional testing. *Journal of Educational computing research*. 24, 29-39.

[21] McNeil, L. (2012). Extending the Compensatory Model of Second Language Reading. *System: An International Journal of Educational Technology and Applied Linguistics, 40* (1), 64-76.

[22] Mohammadi, M. & Barzgaran, M. (2010). Comparability of Computer-based and Paper-based Versions of Writing Section of PET. *The Journal of Applied Linguistics, (3)* 2, 144-165.

[23] O'Malley, J. M. & Chamot, A. V. (1990). *Learning Strategy in Second Language Acquisition.* Cambridge: Cambridge University Press.

[24] Park, H. (2012). *Four English language learners' experiences and strategy use in learning environments of multiliteracies.* Retrieved from University of South Florida, FL.

[25] Park, H. & Kim, D. (2016). English Language Learners' Strategies for Reading Computer-Based Texts at Home and in School. *CALICO Journal, 33* (3), 380-409.

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

[27] Radzi, Amizura Hanadi Mohd, & Aziz, Noor Hashima Abd. (2014). Exploring Content Schemata Influence on L2 Reading: "The Hunted Fox" and "Twelve and Not Stupid". *Advances in Language and Literary Studies, 5* (5), 169-180.

[28] Sawaki, Y. (2001). Comparability of conventional and computerized tests of reading in a second language. *Language Learning & Technology, 5* (2), 38–59.

[29] Smuck, M. (2009). *Reading Is Thinking: Improving Reading Comprehension Using Reading Strategies.* Retrieved from ProQuest Dissertations and Theses.

[30] Sun, L. (2011). *Investigating Chinese English Majors' Use of Reading Strategies: A Study on the Relationship between Reading Strategies and Reading Achievements (Dissertation).* Retrieved from http://hkr.divaportal.org/smash/get/diva2:430259/FULLTEXT01.pdf.

[31] Unzueta, C. H., & Barbetta, P. M. (2012). The effects of computer graphic organizers on the persuasive writing of Hispanic middle school students with specific learning disabilities. *Journal of Special Education Technology, 27 (3),* 15-30.

[32] Worrell, J., Duffy, M., Brady, M., Dukes, C., & Gonzalez-Dehass, A. (2016). Training and Generalization Effects of a Reading Comprehension Learning Strategy on Computer and Paper–Pencil Assessments. *Preventing School Failure: Alternative Education for Children and Youth, 60* (4), 267-277.

[33] Zainal, Z. (2007). Case study as a research method. *Jurnal Kemanusiaan 5* (1), 1-6.

[34] Zhang, D. (2012). Vocabulary and Grammar Knowledge in Second Language Reading Comprehension: A Structural Equation Modeling Study. *Modern Language Journal, 96* (4), 558-575.

[35] Zheng, Y., & L. Cheng. (2008). College English Test (CET) in China. *Language Testing*, 25 (3), 408-417.