

An Empirical Study of the Predictability of Vocabulary Knowledge and Reading Ability on Redesigned SAT Critical Reading and Implication for English Testing System in China

Wei Xu

English Department of Hubei University of Traditional Chinese Medicine
Wuhan, China

e-mail 4174058@qq.com

*Corresponding author Wei Xu

Abstract—This paper mainly discusses the relationship between vocabulary width, lexical depth and critical reading ability for intermediate and high-level proficiency level second language learners in China. By Pearson correlation and linear regress analysis of six types of data, it is found that vocabulary size is unable to predict learners' critical reading proficiency while depth of vocabulary knowledge can effectively predict language learners' critical reading ability and the TEM8 (Test for English Majors in china) has relatively high predictive power of learners' critical reading ability. Text difficulty analysis tool shows that in spite of common test objectives of SAT critical reading and china's TEM8, the SAT critical reading has higher text difficulty and two additional dimensions of test construct: ability to conduct rhetoric analysis and information synthesis.

Keywords—Critical reading ability; redesigned SAT reading; depth of vocabulary knowledge; TEM 8 reading; text analysis tool

I. INTRODUCTION

In English reading, vocabulary has always played a very important role. According to Qian [5], vocabulary knowledge includes two aspects, namely, breadth and depth. In recent years, most researchers believe that the depth of vocabulary knowledge can be used to predict language comprehensive ability and there is a strong correlation between vocabulary breadth and reading ability. On this issue, Lyu [4] is the first in person in China to explore this problem, his research shows a positive correlation between vocabulary size and depth, and vocabulary and language proficiency are positively correlated, but the predictive power of vocabulary for language comprehensive ability is different due to the learners' language proficiency level. Another scholar, Zhang [8] points out that the correlation between the depth of vocabulary and reading is higher than breadth of vocabulary and reading. In other words, the predictive power of depth of vocabulary is better than breadth of vocabulary in reading. In addition, Wang's study [9], from another side also confirmed the conclusion of Zhang and Lyu. His research shows that vocabulary breadth has certain predictive power for CET-4 (College English Test-level 4) and they are moderately correlated, but there is no correlation between CET-6 and vocabulary size, meaning that with

increase of difficulty of texts, vocabulary breadth cannot effectively predict reading comprehensive level. This paper mainly analyzes which key factors can affect critical reading ability of the students through two vocabulary size tests, a vocabulary depth test and results of the standard national tests including TEM (Test for English Majors in China)-level4, the TEM-8. Based on previous researches, we hold a belief that there is a threshold effect when it comes to the effect of vocabulary breadth and depth on reading ability. Therefore, In this study we limit the research subjects to English majors with intermediate and high language proficiency. We select American SAT test as a representative of critical reading because of three reasons: 1) it needs test-takers to find evidence in a passage (or pair of passages) that best supports answer to a previous question or serves as the basis for a reasonable conclusion ;2)it needs the test-takers to identify how authors use evidence such as facts, examples to support their claims and how readings use reasoning to develop ideas and to connect claims and how authors use stylistic or persuasive elements to add power to central claim;3)test-takers need to find a relationship between an informational graphic and the passage it's paired with. [3]

We have three research questions: 1) what is the relationship between English comprehensive ability, reading ability and critical thinking ability for Chinese second language learners? 2) what is the relationship between vocabulary width, lexical depth and critical reading ability for intermediate and high level Chinese second language learners? 3)How can we improve English testing system in china to test critical reading ability of test-takers?

II. RESEARCH PROCESS

A. Research subjects

The subjects are senior English majors at Hubei University of traditional Chinese medicine, and the actual number of participants is 51. They have participated in the TEM-4 in 2015. The VLT, VST, TEM-8 reading tests are conducted during fall semester of year 2016 and spring semester of 2017.

B. Method

The data are mainly collected from five tests, including two vocabulary tests, a vocabulary depth knowledge test, the scores of TEM-4, the reading scores of TEM-8 and scores of redesigned SAT critical reading. The paper adopts the vocabulary size test (VST) which is developed by Nation & Beglar[1] to collect receptive vocabulary; Vocabulary levels test (VLT) designed by Schmitt in 2001[6] is chosen as another vocabulary test and it is a new version after validation of VLT which adopted more scientific Academic Word List designed by Coxhead. VLT has been used to predict the subjects' vocabulary size by measuring single meaning of content words at five frequency levels (2000, 3000, 5000, the university-word level, and 10000). We leave out the 2000 level because of our subjects' language proficiency level. The final test has 120 words with one score for each correct answer. VST has a total of 140 words with one score for each correct answer. For the redesigned sat critical reading part, we select four passages from official guide with 35 test items with one score for each correct answer. For the vocabulary depth knowledge test(VDKT), we compile a Contextual High Frequency Academic Vocabulary Test Table with 30 academic words and give each academic word a corresponding passage. During the test, students are required to read the passage first and figure out the contextual meaning of the academic words.

C. Data collection

The data of VLT, VST tests are carried out in the form of mid-term examination in November 2016 with a total of 120 minutes in which the vocabulary test is 90 minutes, and the reading test is conducted within 30 minutes. The vocabulary depth knowledge test is integrated into the final examination at the end of December 2016, a total of 30 minutes; a SAT critical reading test is conducted on the next week within 70 minutes using online testing system. TEM-8 reading scores are collected after national test results release in June 2017. After all the data are collected, the statistical analysis is carried out by using the statistical package SPSS22.0 and correlation analysis and multiple linear regression analysis.

III. RESULTS

Using SPSS to conduct reliability analysis, the Cronbach's Alpha is 0.75, indicating a high degree of internal consistency between these data for further analysis.

TABLE I. CORRELATION BETWEEN FIVE TESTS AND SAT CRITICAL READING (**P<0.01, * P<0.05)

Variables	1	2	3	4	5
1TEM4	—				
2 vst	0.547**	—			
3 vlt	0.144	0.192	—		
4TEM-8 reading	0.058	0.184	0.297*	—	
5 Vocabulary depth knowledge test	0.376**	0.358**	0.619**	0.223	—
6 SAT critical reading	0.219	0.183	0.514**	0.374**	0.686**

The correlation coefficient between TEM-4, vst and the sat reading scores are below 0.3, and we can rule out those two

variables. The results indicate three variables of vlt and TEM-8 reading and VDKT have significant correlation with sat reading scores. Now that there exists significant correlation among those three variables, does it mean that all of them can effectively predict the critical reading ability? We first use multicollinearity test and find out the tolerance is about 0.952 with 1.051 CI. All the results prove there is no multicollinearity and we can conduct a multiple regression analysis. By gradual multiple regression analysis, we rule out another insignificant variable of VLT. Only the two variables of TEM-8 reading and VDKT are kept afterwards. This confirms our idea that after certain level, the threshold effect kicks in and skills and abilities involved in lower level tests in the TEM 4 and VST(vocabulary breadth test) are not effective anymore in predicting the critical reading abilities.

TABLE II. REGRESSION ANALYSIS

Model	R	R2	R2 Increase
1 Vdkt	0.662	0.438	
2 Tem-8 Reading And Vdkt	0.690	0.476	0.038

The regress analysis shows that in model one when VDKT is put into regression equation, R2 =0.459, meaning deep vocabulary knowledge can explain 45.9% of the changes of critical reading; when TEM-8 reading is entered, R2 increases to 0.502. That is to say, the TEM-8 reading has 5.2% contribution to the changes of critical reading. Then we can build a regression equation to find out the predictive degree. The final equation is this standard curve regression equation: Critical Reading=0.6034*VDKT (vocabulary depth)+0.233*TEM-8reading. The standard regression equation indicates that VDKT (vocabulary depth) can predict critical reading at 60.34% and the TEM-8 reading can predict 23.3% of critical reading scores.

IV. DISCUSSION

Why is there a strong correlation between DVKT and sat Critical Reading? The DVKT requires test-taker to make use of the context to figure dynamic meanings of those 30 academic words after reading 30 short passages. Therefore, this vocabulary depth testing tool requires deep processing of information and comprehensive information summarization and synthesis capabilities. As for the question between high correlation between TEM-8 reading and SAT reading, we conducted a comparative study about text difficulty level of test by using TextEvaluator® designed by ETS[3]. TextEvaluator® text difficulty tool has eight core index: Academic vocabulary and academic word list ratio; Syntactic complexity; Concreteness; Word unfamiliarity; Interactivity style; Narrativity; Cohesion; Argumentation. In following figure, orange color indicates four sat reading article while blue color indicates TEM-8 reading in china. All component scores are reported on a scale that ranges from 1 to 100.Overall text complexity scores are reported on an alternative quantitative scale that ranges from 100 to 2000.The table clearly indicates that the three out of four articles in the US college entrance examination in terms of suitable grade level and text difficulty value are generally higher than TEM-8 reading. If we further

analyze the requirements of those two reading tests, we can find out that SAT critical reading test has 2 additional dimensions of rhetoric analysis and information synthesis different from TEM-8 reading requirements. Those two dimensions reflect the higher critical ability of Bloom taxonomical model, namely, application, analysis, synthesis and evaluation of information. [2].

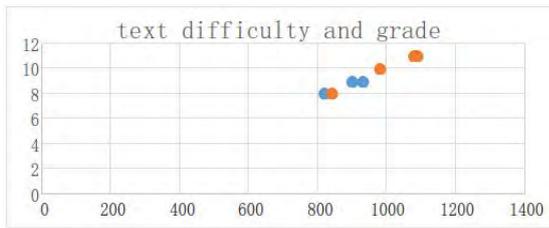


Fig. 1. Analysis of text difficulty and grade of two tests

V. CONCLUSION

What is the relationship between English comprehensive ability, reading ability and critical thinking ability for Chinese second language learners? The English comprehensive ability (namely the integrated ability to listen to news, grammar knowledge, writing an argumentative essay) demonstrated by TEM-4 no long has any predictive power on the SAT critical reading ability. The reason is that when the second language ability is at a limited level, the second language reading will be affected by the language ability, but once it passes the threshold, the reading becomes a complicated process of identifying explicit or implicit central idea and overall textual structure, understanding textual evidences, the tone of academic or historical documents and perspective of writers, finding relationship between an informational graphic and the passages it's paired with or relationship between paired passages. The sat critical reading involves higher level of thinking skills: application, analysis, synthesis and evaluation. [2] What is the relationship between vocabulary width, lexical depth and critical reading ability for intermediate and high level Chinese second language learners? There is no significant correlation between vocabulary breadth and thinking reading ability, but vocabulary depth testing can be used as an important reference indicator for investigating critical reading ability. It has important implication in teaching. The training of lexical depth should be combined with context to improve the ability of L2 learners to understand high frequency polysemy

words and strengthen the ability of deep vocabulary processing. At the same time, the vocabulary depth testing tool (Context High Frequency Academic Vocabulary Test Table) also provides a new research direction for future vocabulary depth testing and English testing system in china.

In 2016, the TEM8 was reformed. The vocabulary of the reading part was maintained at 3000-word count, but the test form was adjusted from the original 20 multiple-choice questions to 14 multiple-choice questions plus 8 short-answer questions. The question type requires candidates to conduct information identification, analysis and reasoning, summarization, and strengthen critical thinking ability in three articles which include both academic and literature readings. The item design not only require the basic information understanding, but also advanced reading skills and ability to analyze the global structure and supporting details of the texts. However, more comprehensive reform is required in the areas of information synthesis and evaluation to set high standards for critical reading ability in china.

REFERENCES

- [1] Beglar, D. A, "Rasch -based validation of the Vocabulary Size Test" in *Language Testing*, Vol. 27, 2010, pp. 101–118.
- [2] Bloom B, Engelhart M, Furst E, et al, "Taxonomy of educational objectives: The classification of educational goals" in *Journal of Nutrition Education*, Vol. 15, 1984, pp. 58-60.
- [3] Board C. Official SAT Study Guide (2016 Edition). COLLEGE BOARD, 2015.
- [4] Changhong Lyu, "The relationship between vocabulary size, depth and language comprehensive ability" in *Foreign Language Teaching and Research*, Vol. 2, 2004, pp. 116-123.
- [5] Qian, D. "Assessing the Roles of Depth and Breadth of Vocabulary Knowledge in Reading Comprehension," in *Canadian Modern Language Review*, Vol. 56, 1999, pp. 282-308.
- [6] Schmitt N, Schmitt D, Clapham C. "Developing and Exploring the Behavior of Two New Versions of the Vocabulary Levels Test," in *Language Testing*, Vol. 18, 2001, pp. 55-88.
- [7] Sheehan K M, Kostin I, Futagi Y, et al. "Generating Automated Text Complexity Classifications That Are Aligned with Targeted Text Complexity Standards," in *Ets Research Report*, 2010, pp. i-44.
- [8] Xuebin Zhang. "Empirical research on relationship between vocabulary knowledge and reading," in *Foreign Language Education*, Vol.1.1, 2006, pp. 38-42.
- [9] Ziying Wang. "Empirical research on predictability of vocabulary knowledge on language ability," in *Foreign Language Teaching*. Vol.1.2, 2014, pp.71-75