Research on the Interactiveness of TEM-4 Writing Tasks

Qinyu Wei[1,a,*]

[1]*Faculty of English Language and Culture, Guangdong University of Foreign Studies, No. 2nd Baiyun Avenue North, Baiyun District, Guangzhou City, Guangdong Province, China*
a. 2901761242@qq.com
corresponding author

Abstract: This study aims to explore the interactiveness of Test for English Majors-Band 4 (TEM-4) writing tasks and tries to give some advice for further improvement. Interactiveness, as defined by Bachman and Palmer, is an important quality in any test. In order to achieve such aims, this study proposes two research questions:1.How interactive is TEM-4's writing? 2.What do candidates think of TEM-4's writing tasks? This study mainly adopts a mixed method which collects data by distributing about 50 questionnaires to English majors online and conducting semi-structured interviews with three participants. Through statistical analysis of the interactiveness of TEM-4 writing tasks from four aspects, this study found that the level of interactiveness of TEM-4 writing task is moderate while the tasks are challenging enough which can truly test students' comprehensive abilities. However, the topics of writing can be more innovative since students feel that they are told in stereotypes and cliches. Therefore, this study broadens the application field of the theory of interactiveness and applies it to TEM-4 writing tasks in China. Furthermore, it also can provide helpful implications for test makers.

Keywords: TEM-4, writing tasks, interactiveness

1. Introduction

The Test for English Majors-Band 4 (TEM-4) is of great significance for English majors which is the standard for evaluating their English proficiency. It is widely recognized in China and is one of the most common certificates for English majors to prove their comprehensive English ability. Therefore, the quality of the TEM-4 is so essential to candidates that it is necessary to conduct language assessment.

Language assessment refers to making inferences about a person's language knowledge and related abilities [1]. In assessing learners, we should figure out how learners acquire languages; what evidence that can demonstrate language knowledge of learners and their corresponding abilities; what difficulties they have and how to improve their abilities. Two kinds of data are often collected through language assessments. The first is about assessing the language learning progress, which is known as educational assessment.The second refers to assessing information to see whether a person's language ability meets certain standard, which is called proficiency assessment. Since the language learning is not emphasized, this kind of assessment, most of which are formal tests, is often administered by national or international agencies. Since learning processes are not essential in proficiency assessment which centers on a person's present functionality, it is quite different from educational assessment. To

be specific, the proficiency assessment aims to see what the assessee can complete by using the language at present without involving the process of language acquirement or time they spend on it. The final goal of this assessment is judging whether the assessee's language ability satisfies some requirements.

Therefore, the designer of tests should clarify what knowledge and skills will be assessed and how these elements are assessed. According to the above discussion, TEM-4 is kind of proficiency assessment in China.

TEM-4 is one of the most authoritative English proficiency tests for English majors in China or more specifically, the tests are for English majors in their fourth term as specified in the National College English Teaching Syllabus for English Majors [2]. After the reform, the new writing tasks in TEM-4 has aroused interest of many researchers most of whom used usefulness to evaluate the quality of the task. The usefulness is an essential quality in a test [3]. A mode of test usefulness including six test qualities was proposed by Palmer, which are reliability, construct validity, authenticity, interactiveness, impact, and practicality. This mode is comprehensive for test assessment, and based on these test qualities, many studies about TEM-4 have been conducted [4-6]. However, although there are many studies about validity, reliability and washback in TEM-4, the study on interactiveness is few.Therefore, this essay attempts to study the interactiveness of TEM-4 and tries to give some advice for further improvement.

There are five chapters in this thesis. The introduction, which is the first chapter, provides a brief overview of the background of the study, its purpose and significance, and the structure of the essay. The second chapter is a review of the literature. The definition of terms, relevant studies, and the theoretical foundations are covered in this chapter. The third chapter provides an overview of the entire research process, including data collecting and analysis, as well as the research object, research questions, research methods, and research tools. The results and discussion section are included in the fourth chapter. The last chapter finishes by summarizing the entire essay while providing some suggestions for how to make TEM-4 writing tasks more interactive.

2. Literature Review

Definitions of terms, the theoretical foundations, related studies, and the purpose and significance of the research are all included in this part.

2.1. Interactiveness

The interactiveness is an important quality in any test and is defined as "the extent and type of involvement of the test taker's individual characteristics in accomplishing a test task" [3]. The individual characteristics which are closely related to language assessment mainly consist of language ability including language knowledge and metacognitive strategies, topical knowledge, and affective schemata.

Language ability is the ability that combines language knowledge and metacognitive strategies to create and explain discourse. When using metacognitive procedures to create and interpret discourse in language use, language knowledge can be seen as a domain of information in memory. Strategic competence consists of a collection of metacognitive elements or strategies, which can be seen as higher-order executive functions that perform cognitive management functions in other cognitive tasks and language use.

It is possible to regard topical knowledge as knowledge structures in long-term memory (also known as knowledge schemata or real-world knowledge). Topical knowledge of individuals should be considered when describing language usage because it gives them the knowledge base necessary to use language concerning their world, which is involved in all language use.

Topical knowledge's emotional or affective correlates are known as affective schemata. These affective schemata serve as the framework on which language users evaluate the features of the language use task and its context in terms of previous emotional experiences in related circumstances, whether consciously or unconsciously.

In summary, interactiveness is affected by the subject's language ability, topical knowledge, and affective schemata, as well as the degree to which each of these factors influences interactiveness.

Therefore, language ability, topical knowledge, and affective schemata should all be considered when analyzing the interactiveness of TEM-4 writing tasks. After the reform, the writing of TEM-4 requires candidates to have further thinking and give comments based on the understanding and induction of the reading text. Moreover, this comprehensive writing task aligns with students' reading and writing patterns in daily life, while substituting material composition for proposition composition can lessen students' likelihood of copying the "stereotyped writing" template sentence pattern. Candidates have more freedom to express their emotions. As a result, after the reform, writing should be highly interactive theoretically [7]. Although there are many studies on TEM-4's validity, reliability, and washback, there are relatively few studies evaluating its interactiveness [8,9]. Then this essay will investigate and examine how interactive the TEM-4 writing task is.

2.2. TEM-4 Writing Tasks

TEM-4 aims to measure the English proficiency of English majors in China. In 2016, after the reform, in writing part, note writing in TEM-4 was canceled while it added reading and required students to summarize the given materials. The new kinds of questions mainly examine students' information extraction, summarizing, reading comprehension, and paraphrasing. The new assessment criteria for writing test are as follows [10]:

Table 1: Writing assessment criteria.

Rating	Score Level	Elaboration	Text Organization	Language Organization
5	18-20	8.5-10	2.5-3	6.5-7
4	15-17	7.5-8	2-2.5	5.5-6
3	12-14	6-7	1.5-2	4.5-5
2	8-11	5-5.5	1-1.5	3.5-4
1	1-7	1-5	0.5-1	1-3
0	0	0	0	0

Generally speaking, writing evaluation mainly focuses on content, logic, organization and expression. After the reform, students are required to express their own opinions and ideas in accordance with the viewpoints and facts provided in the materials. This procedure will further test the students' capacity for integrating information. Therefore, Students can thus employ the knowledge they obtained from reading and studying in daily circumstances such as the articles they have learned and the deeds they have heard, and so on. This demonstrates that these topics possess the interactiveness that the usefulness framework has suggested.

3. Methodology

This part provides a detailed introduction to the research subjects, methods of investigation, research procedure, research questions and ethical considerations. All these are to provide a guarantee for the analysis of questionnaire survey results and the interviews. They aim to better evaluate the

interactiveness of TEM-4's writing by combining quantitative and qualitative methods.

3.1.   Research Questions

The purpose of this study is to evaluate the quality of TEM-4's writing based on Bachman and Palmer's model of interactiveness and, at the same time, make some suggestions to the improvement of the writing tasks' quality. According to Bachman and Palmer, interactiveness consists of four elements. That is to say, evaluating the quality of a test paper is achieved by carefully evaluating its interactiveness in four aspects. Therefore, this essay puts forward the following two research questions, trying to give an objective evaluation:

1. How interactive is TEM-4's writing?
2. What do candidates think of TEM-4's writing tasks?

3.2.   Research Subjects

In the questionnaire survey, the research object of this topic is composed of 26 girls and 24 boys aged 21~22. Participants are all English majors from different regions of China who have taken the TEM-4 exam, representing different learning levels. In the semi-structured interview, the participants are three girls randomly selected from all of the participants in the questionnaire survey.

3.3.   Research Methods

(1) Examination paper

This study uses the examination papers of TEM-4 in 2021 and 2022.

(2) Questionnaire survey and semi-structured interviews

This study used an online questionnaire survey and semi-structured interviews to collect data. And the results from the questionnaire survey and interview will answer the first question and the data collected from interview will explain for the second question. In order to make every research object understand the questions of the questionnaire clearly, this questionnaire is in Chinese. The questionnaire mainly covers three aspects that affect interactiveness: language ability, topical knowledge, and affective schemata. Therefore, this questionnaire investigates the interactiveness of the test paper from these three dimensions. Questions 1~5 examines the language ability, question 6 examines the topical knowledge, and 7~9 examines the affective schemata. The topics are all based on Bachman & Palmer's theory framework [3]. The questionnaire contains nine multiple-choice questions, and the choice of each question is calculated according to the score, from low to high, ranging from 1 to 5 points. The higher the score, the higher the approval. If the score is 5, it represents the strongest support attitude; if it is 3, it will remain neutral; If the score is 1, it means it is not supported.

In semi-structured interviews, three students are randomly selected to answer some questions. For example: "What do you think of TEM-4 writing task?" "What suggestions can you provide for the TEM-4's writing task?"

3.4.   Experimental Process

The analysis of interactiveness in this study adopts a combination of quantitative and qualitative analysis. First, a questionnaire survey is conducted among English majors, followed by semi-structured interviews. The specific operation steps are as follows: 1. Based on the questions of evaluating the interactiveness of test papers put forward by Bachman and Palmer in the book Language Testing in Practice, the new questions are proposed for TEM-4 writing task. 2. Collecting data from candidates online through questionnaires. 3. Interviewing three students randomly so as to further analyze the results of the questionnaires.

3.5. Data Collection

In this study, a questionnaire survey was conducted among junior and senior English majors by distributing and collecting data online. The time of this survey begins at 17: 00 pm on May 1, 2023, and ends on May 7. Seventy-five questionnaires were collected, of which 50 were valid, and 25 were invalid. On May 8, the author randomly selected three of the 50 valid questionnaires and conducted a semi-structured interview with the respondents. The author took accurate notes during the interview with the students, taped them simultaneously, and later transcribed the content.

3.6. Ethical Consideration

The participants were aware of the study's objective, and students' permission was secured before conducting the experiment.

4. The Research Results and Discussion

In this part, based on Bachman and Palmer's criteria for evaluating interactiveness, the author will evaluate the interactiveness of TEM-4 writing task through analyzing the data collected from the questionnaire survey and semi-interviews. This essay will evaluate interactiveness in four aspects:
(1) Is the language knowledge of the TEM-4 writing task suitable for candidates?
(2) Can candidates use strategic competence to help them perform better in the exam?
(3) Is the writing task's topical knowledge suitable for candidates?
(4) Can the writing task help or impede candidates' ability to use language?
Based on Oxford's Strategy Inventory for Language Learning (SILL), the study employed the following standards to evaluate the interactiveness of the language knowledge, strategic competence, topical knowledge and affective schemata [11].

Table 2: Standards to evaluate the interactiveness.

Score of involvements	Level of interactiveness
5.0>score>=3.5	High
3.5>score>=2.5	Moderate
2.5>score>=1	Low

4.1. The Language Knowledge of the Writing Task

TEM-4 writing task mainly tests the students' comprehensive ability, including vocabulary, grammar, and reading. Therefore, to complete the task, candidates must have these basic language skills, not just vocabulary. Given the dimension of language knowledge in interactiveness, the author conducted a questionnaire on candidates, and the results are as follows:

Table 3: Candidates' attitude towards the language knowledge involved in the test paper.

	1	2	3	4	5	Mean score
Q1	0%	30%	34%	26%	10%	3.16
Q2	2%	18%	46%	30%	4%	3.16

From Table 3, we can draw a conclusion that in the first question, 34% of the candidates think that they sometimes relate to their previous knowledge when answering questions. And 26% of the candidates choose "often," while 10% of the candidates choose "always," and 30% of the candidates

choose "occasionally", and no candidate chooses "never". In the second question, 46% of the candidates think that they sometimes use a variety of language knowledge during the writing test, 30% choose "often," 4% choose "always," 18% choose occasionally, and only 2% choose "never". In the semi-interview, all of the three students agree that the writing task tests various language abilities and the task is very comprehensive. From the above analysis and survey data, the language knowledge of TEM-4 writing task is relatively suitable for candidates, and it involves a wide range of language abilities of candidates.

4.2.　The Strategic Competence of Candidates

Since the summary part was added and the note writing was deleted after the reform in 2016, the difficulty has increased. What kind of strategies to use to answer questions varies from student to student. This study did a survey based on several common writing strategies, and the results are shown in the following table:

Table 4: Candidates' use of writing strategies.

	1	2	3	4	5	Mean score
Q3	2%	32%	28%	30%	8%	3.1
Q4	4%	22%	38%	24%	12%	3.18
Q5	6%	22%	36%	28%	8%	3.1

From the survey results, it is evident that only about 30% of the candidates often set a target score for themselves before taking the exam, making a draft or write the structure before formal writing, checking and revising it after finishing the writing. Another 30% of the candidates sometimes use these strategies, and the rest occasionally or even do not use any strategies. This shows that the candidates can generally realize that they need to use metacognitive strategies to help them perform better in the exam.

4.3.　The Writing Task's Topical Knowledge

The topics examined in the TEM-4 writing tasks are relatively broad, such as topics about the relationship between wisdom and loneliness and environmental protection in 2021 and 2022, which are also common topics for candidates.

Aiming at testing the dimension of topical knowledge in interactiveness, the author conducted a questionnaire survey on candidates, and the results are shown in the following table:

Table 5: Candidates' familiarity with the topic knowledge involved in the test.

	1	2	3	4	5	Mean score
Q6	4%	16%	44%	32%	4%	3.16

From Table 5, we can see that 44% of the candidates choose that they sometimes use their familiar experiences to answer questions , 32% choose often, 4% choose always, and 20% choose never and occasionally. From this, most candidates will be more or less willing to use their familiar experiences as materials when writing. From the interview, both student A and C pointed out that in TEM-4 writing test, the topics are not very creative, and students are quite familiar with the topics. Therefore, the writing task's topical knowledge is relatively suitable for candidates.

4.4. The Affective Schemata in the Test

Familiar topics or topics that candidates like will stimulate candidates' writing inspiration, while obscure content and unfamiliar topics will hinder candidates' performance; Moreover, the allocation of examination time will also affect the performance of candidates.

Aiming at testing the dimension of affective schemata in interactiveness, the author conducted a questionnaire survey on candidates, and the results are shown in the following table:

Table 6: Candidates' affective schemata in the test.

	1	2	3	4	5	Mean score
Q7	2%	16%	56%	22%	4%	3.1
Q8	8%	24%	34%	30%	4%	2.98
Q9	2%	20%	44%	34%	0%	3.1

From Table 6, for topics involved in these two exams, half of the candidates do not show any attitude about the topics involved in the exams, and it is possible that they are not sensitive to the topic knowledge. Among the other half of the candidates, 26% choose "like", which is higher than choosing "dislike". From the interview, both student A and B feel that the topic fits students a lot since it requires students to think deeply. Therefore, the affective schemata of TEM-4 writing task is relatively suitable for candidates. For the question types after reform, the number of students who choose "like" and the number of students who choose "dislike" are almost the same, and the remaining 30% feel that the questions are general and have no influence on them. From the interview, both student B and C claim that the questions are not very easy while student C even feels they are challenging after the reform. Therefore, on the whole, the question types in TEM-4 writing tasks are relatively acceptable to most candidates.

From the table, although 22% of the candidates feel that time is insufficient, most candidates feel that the time given for writing is enough. Moreover, from the interview, the student C who chose "insufficient" declares that she will spare some time from the time for writing for previous questions. Therefore, the time allocation for writing is reasonable in some degree. Therefore, on the whole, TEM-4 writing task will not hinder the candidates from exerting themselves in the dimension of affective schemata.

According to the above analysis of the interaction of English writing in TEM-4, we can know that candidates generally use metacognitive strategies to answer questions, and the TEM-4 writing task will not hinder the candidate from answering questions. At the same time, the language knowledge that the test examines has already been learned in students' usual study. Therefore, in general, the level of interactiveness of TEM-4's writing test is moderate.

5. Conclusion

This part concludes the main findings of the study, some suggestions for TEM-4 writing tasks and the limitations of this study.

5.1. The Main Findings of Research

The level of interactiveness of TEM-4 writing task is moderate.The exam's topic of study is familiar to the candidates and closely related to their daily lives. It can arouse the affective schemata of candidates and make them try to extract the knowledge stored in their minds to respond. Moreover, the tasks are not very easy which can truly test students' comprehensive abilities.

5.2. Suggestions for TEM-4 Writing Task

Since almost all of the topics involved in the writing tasks are ones that the students are already familiar with, new topics relevant to everyday life may be encouraged to add.

5.3. Limitations of Research

Although this study has tried to analyze it as comprehensively as possible, there are still some shortcomings. First, it is inevitable that the analysis may be influenced by subjective feelings, which is not objective enough. Second, the sample of the questionnaire survey is 50 for students. The sample is not large enough and can only represent the opinions of some students. Lastly, this study only analyzes TEM-4 writing tasks in 2021 and 2022, which may be not sufficient.

References

[1] Green, A. (2020). Exploring language assessment and testing: Language in action. Routledge.
[2] NACFLT. (2000). Syllabus for university English language teaching. Shanghai Foreign Language Education Press & Foreign Language Teaching and Research Press.
[3] Bachman, L. F., & Palmer, A. S. (1996). Language testing in practice: Designing and developing useful language tests. Oxford University Press.
[4] Guo, L., & Chen, Y. (2007). A comparative study of CET-4 and TEM-4 in China. Asian Englishes, 10(1), 90-102.
[5] Tian, X. H. (2015). A study on the reliability and validity of TEM-4. Teaching of Forestry Region, 1, 42-43.
[6] Treffers-Daller, J., & Huang, J. (2020). Measuring reading and vocabulary with the Test for English Majors Band 4: A concurrent validity study.
[7] Liu , M. (2019). An analysis on the reform of writing task in TEM-4. Modern Communication, 19, 11-12.
[8] Bing, H., Hui, T., & Lirui, L. (2012). An analysis of the washback effect of TEM-4 on teaching of English majors– taking Beijing Sport University as an example. Multilingual Theory and Practice in Applied Linguistics, 23.
[9] Xu, Q., & Liu, J. (2018). A study on the washback effects of the Test for English Majors (TEM). Springer.
[10] Mao, N. (2020). Analysis of writing skills in TEM-4. Think Tank Era, 09, 182-183.
[11] Oxford, R. (1990). Language learning strategies: What every teacher should know. Newbury House.