

A Study on Content Validity of Reading Comprehension in TEM-8 (2016-2021)

Chen Xingyu

Guangzhou College of Applied Science and Technology, Guangzhou, 526000, Guangdong, China

Abstract: *TEM-8 is an authoritative test to judge the professional level of English majors, in which reading comprehension occupies an important proportion. Based on Bachman & Palmer's test task characteristics framework and combined with the requirements of teaching syllabus for English majors and the revised TEM-8 test syllabus in 2015, this paper, from the perspectives of language input characteristics and expected response characteristics, takes TEM-8 reading comprehension passages in the past five years as the research object to analyze the content validity of reading comprehension with the methods of qualitative and quantitative analysis. The results reveal that TEM-8 reading comprehension test in 2016-2021 has relatively high content validity on the whole, but there is space for further improvement in the length of reading text, the selection of text materials, the balance of text difficulty and the diversity of response types.*

Keywords: *content validity; reading comprehension; language test; TEM-8*

1. Introduction

As the highest-level nationwide test for English majors in China, TEM-8 (Test For English Major-Band 8), undoubtedly, is an important EFL (English as a foreign language) test in China. It is administrated by the National Advisory Commission on Foreign Language Teaching in Higher Education (NACFLT) aiming to measure the actual English proficiency of university undergraduate English majors in accordance with the National College English Teaching Syllabus for English Majors. Furthermore, it has gradually become a stepping stone for English majors to obtain opportunities of employment and promotion.

Nevertheless, it is universally acknowledged that the TEM-8 is the hardest English language test in terms of measuring the integrative English competence of test takers. It measures a variety of English skills of test takers, including listening, reading, language usage, translation and writing. Among different parts of TEM-8, reading comprehension always accounts for the largest proportion, especially after the revision of TEM-8 test syllabus in 2015, the proportion rose from 20% to 30%, which demonstrates the critical role of reading comprehension in TEM-8. However, whether the TEM-8 papers in recent years have tested the desired effect as expected remains to be further studied.

Therefore, this paper uses test task characteristics as a theoretical framework to study the reading comprehension in TEM-8 from 2016 to 2021, focusing on whether TEM-8 reading comprehension has high content validity and what problems remain to be improved in content validity. This paper aims to make a scientific interpretation of the content validity of TEM-8 reading comprehension, and also hopes to enlighten and guide the validity research of language test.

2. Content validity and theoretical framework

When it comes to validity, Hughes (2003: 26) pointed out that “a test is said to be valid if it measures accurately what it is intended to measure”^[1]. There are three forms as the evidence of validity and Hughes (2003) named them as content validity, criterion-related validity and construct validity^[1]. Of which, the first form, content validity, is considered as the target to study in the reading comprehension of TEM-8 because this paper mainly focuses on the validity of the test text.

According to Popham (2002: 52), content validity refers to ‘the extent to which an assessment procedure adequately represents the content of the assessment domain being sampled’^[2]. Hughes (2003) argued that in order to judge whether or not a test has content validity, a specification of the skills or structures could be necessary^[1]. All in all, the guarantee of content validity mainly depends on the

coverage analysis of test questions before the test, and the verification is based on the syllabus^[3].

When referring to the theoretical framework, Bachman and Palmer proposed the framework of test task characteristics in the 1980s. It contains the representative factors of the language skills and language structures that it intends to test. In addition, both the teaching syllabus for English majors and the test syllabus of TEM-8 specify the test requirements for TEM-8. Therefore, take the framework of Bachman and Palmer as the cornerstone, the research of this paper combines with the teaching and test syllabuses to analyze the content validity of reading comprehension sections in TEM-8 from the perspective of reading task characteristics, including input characteristics (text length, new words, topic, genre and readability) and the expected response characteristics (reading skills coverage and response type).

3. Analysis and discussion of content validity of tem-8 reading comprehension

In this paper, the research objects are 5 test papers covering 15 passages in TEM-8 reading comprehension from 2016 to 2021 (TEM-8 in 2020 was canceled due to the 2019 Novel Coronavirus). It will utilize qualitative and quantitative research methods to analyze TEM-8 reading comprehension in recent five years based on the revised theoretical framework mentioned above to test how high its content validity is.

3.1. Input characteristics

Input characteristics are mainly reflected in five aspects: text length, reading speed, text difficulty, topic and genre.

3.1.1. Text length and reading speed

Text length here refers specifically to the length of each passage. According to the point of view of Bachman (1990), the longer the reading passage, the greater the impact on its discussion characteristics^[4]. What's more, the reading passages of TEM-8 should be about 3,000 words in terms of the revised test syllabus in 2015. As a consequence, words counting calculated by the Microsoft Word can be shown clearly in Table 1.

Table 1: Text length and reading speed of passages for TEM-8 reading comprehension

Year	Passage words			Total words of passage	Reading Speed (words/minute)
	Passage A	Passage B	Passage C		
2016	996	1014	968	2978	119.12
2017	738	1235	845	2818	112.72
2018	835	859	879	2573	102.92
2019	1026	992	919	2937	117.48
2021	735	1258	1035	3028	121.12

As it shown in Table 1, the text length of TEM-8 reading passages from 2016-2021 is from 2,573~3,028 words, all of which are about 3,000 words, in line with the requirements of test syllabus (2015). However, the text length in 2018 is slightly smaller, and the rest is relatively average distribution.

After the revision of TEM-8 test syllabus in 2015, the reading time was extended from 30 minutes to 45 minutes. Considering that the test takers also need a certain amount of time to read the question items and think about the answers, so the proportion of time needed to read the passages and answer the questions should be 1:0.75^[5]. Furthermore, parts of the reading comprehension questions have been changed from multiple-choice questions to short answer questions, which virtually mounts the time for test takers to answer questions. As a result, the author comes to the conclusion that the test taker needs roughly 25 minutes to read the passages and 20 minutes to answer the questions based on the comprehensive consideration.

When it comes to the reading speed, the test syllabus (2015) specifies the reading speed should be about 150 words per minute. As shown in Table 1, from 2016 to 2021, the reading speed is 103-121 words per minute, far below the reading speed required by the syllabus, which generally lowered the content validity of TEM-8 reading comprehension in the five years. In other words, the reading time has increased while the text length has decreased, which is the main reason why the reading speed is not up to the standard of the test syllabus.

3.1.2. Topic and genre

As for topic area and genre, Weir (1993: 67) points out “the topic should be selected from a suitable genre, at an appropriate level of specificity, and should not be culturally biased or favor one section of the test population” and he also emphasizes that at a high level of proficiency, appropriate text genres should be chosen for the test takers, such as narration, description, and argumentation, etc^[6]. Additionally, test syllabus (2015) stipulates the following principles for the selection of reading materials in TEM-8: (1) a wide range of topics, including society, science and technology, culture, economics, literature, linguistics, biography, etc.; (2) various genres, including narration, description, exposition, argumentation and so on.

As can be seen from Table 2, topics of reading materials from 2016 to 2021 mainly focus on society and literature, each accounting for 33.3% of the total, followed by the topic of culture, accounting for 26.7%, and finally, only a small amount of topic are related to science and technology (6.7%). However, economics, linguistics and biography were absent for five years.

Table 2: Topic areas of passages for TEM-8 reading comprehension

Year	Society	Science and Technology	Culture	Economics	Literature	Linguistics	Biography
2016	1	1	0	0	1	0	0
2017	1	0	1	0	1	0	0
2018	1	0	1	0	1	0	0
2019	1	0	1	0	1	0	0
2021	1	0	1	0	1	0	0
Proportion	33.3%	6.7%	26.7%	0	33.3%	0	0

As shown in Table 3, in the reading materials for TEM-8 from 2016 to 2021, the four genres stipulated in the test syllabus (2015) are all involved in the selection of materials in these five years, but the proportion of descriptive articles is relatively small. According to the principle of genre diversification, it is suggested that other genres should also be considered, such as advertisements, specifications, diagrams or charts and so on. To sum up, after the revision of the test syllabus in 2015, the breadth of the topics for TEM-8 reading comprehension is not satisfactory, but the genres can basically meet the requirements of diversification.

Table 3: Genre distribution of passages for TEM-8 reading comprehension

Year	Narration	Description	Exposition	Argumentation	Others
2016	1	0	1	1	0
2017	1	1	0	1	0
2018	1	0	1	1	0
2019	1	0	1	1	0
2021	1	0	1	1	0
Proportion	33.3%	6.7%	26.7%	33.3%	0

3.1.3. Text difficulty

Generally, the study of text difficulty mainly refers to the research of readability at home and abroad. Li Shaoshan (2000: 1) states that “readability means the ease with which a text is read and understood”^[7]. Meanwhile, he concludes that lexical (especially word frequency) and syntactic difficulty have high predictive ability for text readability^[7]. Therefore, it is considered to use the new word rate and the data measured by Flesch readability formula as the basis to judge the difficulty of the passages for TEM-8 reading comprehension.

Table 4: New words of passages for TEM-8 reading comprehension

Year	New words	Total words of the passages	Rate of new words (%)
2016	17	2978	0.57
2017	18	2818	0.64
2018	9	2573	0.35
2019	15	2937	0.51
2021	10	3028	0.33
Average	13.8	2866.8	0.48

As can be seen from Table 4, the rate of new words in TEM-8 reading materials from 2016 to 2021 is 0.33%-0.64%, basically controlled below 1%, in line with the requirement of the teaching syllabus for English majors, that is, the new words of passages should not exceed 3%.

As for readability, it is calculated carefully with the help of the Software Readability Analyzer 1.0. Its data and the distribution can be presented in the following tables (Table 5 & 6).

Table 5: Readability of passages for TEM-8 reading comprehension

Year	Passage A	Passage B	Passage C	Average
2016	60.3	38.3	45.8	48.13
2017	54.7	56.9	36.3	49.30
2018	43.9	51.9	71.5	55.77
2019	75.6	55.2	61.1	63.97
2021	61.4	74.1	39.2	58.23

Table 6: Readability distribution of passages for TEM-8 reading comprehension

Year	30-50 (difficult)	50-60 (fairly difficult)	60-70 (standard)	70-80 (fairly easy)	80-90 (easy)	Total
2016	66.7	0	33.3	0	0	100
2017	33.3	66.7	0	0	0	100
2018	33.3	33.3	0	33.3	0	100
2019	0	33.3	33.3	33.3	0	100
2021	33.3	0	33.3	33.3	0	100
Average	33.3	26.7	20	20	0	100

From Table 5 and Table 6, some conclusions can be drawn as follows.

In general, the higher the readability score is, the simpler the text is. Therefore, according to Table 5 and Table 6, only the average Reading Ease score of the reading comprehension passages in 2019 reaches “standard” (63.97), and the rest scores are between 48 and 58, that is to say, the reading difficulty of the passages is between “difficult” and “fairly difficult”. Hence, it can be concluded that the majority of reading passages from 2016 to 2021 are at the level of “difficult” and “fairly difficult”, with a total percentage of 60% (33.3%+26.7%). According to the teaching syllabus, English juniors and seniors must be able to read editorials and book reviews from ordinary English newspapers and magazines, as well as difficult historical biographies and literary works. Consequently, most passages in reading comprehension meet the requirement of teaching syllabus for English majors, and thus, the conclusion can be drawn that the content validity is relatively quite good.

Yet, in terms of specific scores of Reading Ease, TEM-8 reading comprehension readability in 2018 is the most special, with the statistics from low to high, being 43.9, 51.9 and 71.5 respectively. Psychologically speaking, reading passages should be read from easy to difficult in the test, and the same suggestion has been put forward by Chen Shifei (2014: 36), “reading passages in the range of reasonable difficulty levels need to be arranged orderly from easy to difficult with the purpose of bringing test takers’ ability in working order”^[8]. As a consequence, in order to further the content validity of TEM-8 reading comprehension, the author suggests that the difficulty of the three passages in TEM-8 reading comprehension should be increased in a gradual way and kept at a stable state as far as possible.

On the whole, it can be seen clearly from the above tables that the new word rate from 2016 to 2021 shows a fluctuating decline while the text readability shows a fluctuating rise, which indicates that the difficulty of TEM-8 passages declines after the revision of test syllabus in 2015, but it is still within the scope stipulated by the test syllabus, and on the other hand, it verifies that there is a relatively high content validity in the difficulty of texts in these five years.

3.2. Expected response characteristics

To test candidates’ reading ability is actually to test their mastery of reading skills. In the framework of test task characteristics, the characteristics of expected response are mainly centered on the reading skills coverage and response type.

3.2.1. Reading skills coverage

In this section, it is necessary to determine the types of reading skills coverage. The author integrates Davis' classifications (cited in Alderson, 2000)^[9] with Zou Shen's (1997) specifications^[10] of reading comprehension to form a revised reading skills coverage consisting of seven kinds, namely:

- a) Distinguishing important facts or details
- b) Identifying a writer's technique
- c) Drawing inferences about the meaning of a word or phrase in context
- d) Drawing inferences from the content
- e) Understanding writer's purpose, attitude to audience and to topic
- f) Determining the main ideas
- g) Understanding discourse structures

Thus, based on the newly formed reading skills coverage mentioned above, the statistics can be presented as follow (see Table 7).

Table 7: Reading skills coverage of passages for TEM-8 reading comprehension

Year	(a)	(b)	(c)	(d)	(e)	(f)	(g)
2016	5	1	2	10	1	3	0
2017	5	0	3	12	1	1	0
2018	6	1	1	11	2	1	0
2019	8	1	5	5	1	1	1
2021	9	1	3	4	2	1	1
Total	33	4	14	43	7	7	2
Proportion	30.0%	3.6%	12.7%	39.1%	6.4%	6.4%	1.8%

First of all, the coverage of reading skills needs to be widespread. As shown in Table 7, TEM-8 reading comprehension in the past two years (2019 and 2021) examined all reading skills of test takers, but one or two skills were not examined in the years of 2016-2018. Secondly, among all the skills examined, item (d) and item (a) rank the first and second respectively, accounting for nearly 70% of the total, while other reading skills are less examined. Finally, it can be summarized that TEM-8 reading comprehension tests from 2016 to 2021 basically assess all the reading skills stipulated in the test syllabus (2015) and achieve relatively high content validity. On the other hand, however, its data also reflect the unbalanced distribution of reading skills to some extent.

3.2.2. Response type

After the revision of test syllabus in 2015, besides multiple-choice question, short answer question is added in the question types of TEM-8 reading comprehension. Specifically, there are 14 multiple-choice questions and 8 short answer questions in reading comprehension, accounting for 47% and 53% of the reading comprehension scores respectively. That is to say, although the numbers of short answer questions are less than multiple-choice questions, their scores are higher, which proves that TEM-8 reading comprehension emphasizes more on examinees' ability of text understanding and language organization after the revision of test syllabus in 2015.

According to Bachman and Palmer (1996), response types consists of three aspects, that is, response of limited production, response of selected and response of extended production^[11]. Moreover, the test syllabus (2015) also makes it clear that TEM-8 aims to test students' comprehensive ability to use English language. As a result, in order to improve the content validity of test, it is essential to enhance the comprehensiveness of TEM-8 reading tasks from the perspective of its test design. Which means, in addition to multiple-choice questions and short answer questions, more types of response should be adopted to strengthen the degree of participation of test takers, such as matching, information transfer, sentence completion, diagram/flowchart/table completion, summary writing and so on.

4. Conclusion

This paper analyzes the content validity of reading comprehension in TEM-8 from 2016-2021 on the basis of the revised assessment framework, mainly focus on two characteristics, namely the language

input and expected response.

The analysis of input characteristics is carried out from the following five points: text length, reading speed, topic area, genre and readability, and the expected response characteristics are represented by reading skills coverage and response types.

Results analysis explicates that TEM-8 reading comprehension tests have relatively high content validity in recent five years from the aspects of language input and expected response. Because of its scientific, rigorous and standardized test design, it basically conforms to the requirements and provisions of TEM-8 test syllabus and English teaching syllabus for English majors. First of all, the difficulty of reading comprehension passages conforms to the requirements of the test syllabus. Secondly, the reading materials have a wide range of topics, involving most of them stipulated in the test syllabus; besides, there are a variety of genres, and the four main genres are all present in the reading materials. Finally, the response type setting is also in accordance with the test syllabus so as to test the reading skills of candidates from different angles.

Nonetheless, TEM-8 reading comprehension test still has space for improvement. In terms of the text length, it can be paid more attention to the length of each text to ensure the balanced distribution of the length. And the selection of text materials can be more close to life, to make sure a wide range of topics and genre diversification. It can be considered to select advertisements, specifications and other aspects of the genre, and can attempt to choose economic and historical articles. In the field of readability, attention should be paid to the balance of text difficulty between each discourse. In other words, control it within a reasonable range and guarantee that there is no much difference in the degree of difficulty. As for the expected response, response types also should be designed to be more flexible and diverse, in consequence, the types of response can be appropriately changed or increased, such as matching, sentence completion, summary writing, etc.

This paper aims to investigate the content validity of TEM-8 reading comprehension in recent five years, whereas does not research the content validity of the whole test paper, so there are some certain defects in this study. Based on this limitation, the author expects that more studies on validity of TEM-8 will be carried out in the future.

References

- [1] Hughes, A. *Testing for Language Teachers*[M]. Cambridge: Cambridge University Press, 2003.
- [2] Popham, W. *Classroom Assessment: What Teachers need to know*[M]. Boston: Allyn and Bacon, 2002.
- [3] Li Xiaojun. *The Science and Art of Language Testing*[M]. Changsha: Hunan Education Press, 2001.
- [4] Bachman, Lyle. F. *Fundamental Considerations in Language Testing*[M]. Oxford: Oxford University Press, 1990.
- [5] Yang Huizhong & Weir, C. J. *Validation Study of the National College English Test*[M]. Shanghai: Shanghai Foreign Language Education Press, 1998.
- [6] Weir, C. J. *Understanding and Developing Language Tests*[M]. New York: Prentice Hall, 1993.
- [7] Li Shaoshan. *A Survey of Studies on Readability*[J]. *Journal of PLA University of Foreign Languages*, 2000 (04): 1-5.
- [8] Chen Shifei. *A Content Validity Study of TEM-8 Reading Comprehension From 2009 to 2013*[D]. Guilin: Guangxi Normal University, 2014.
- [9] Alderson, J. C. *Assessing Reading*[M]. Cambridge: Cambridge University Press, 2000.
- [10] Zou Shen. *TEM Validation Study*[M]. Shanghai: Shanghai Foreign Language Education Press, 1997.
- [11] Bachman, Lyle. F. & Palmer, A. S. *Language Testing in Practice: designing and developing useful language tests*[M]. Oxford: Oxford University Press, 1996.