

A Study of Content Validity on Grammar Gap-filling Tests in NMET II*

Ming-Hao Jin · Mei-Xian Zheng**
(Yanbian University)

Jin, Ming-Hao & Zheng, Mei-Xian. (2020). A study of content validity on grammar gap-filling tests in NMET II. *The Linguistic Association of Korea Journal*, 28(3), 71–81. Based on Bachman and Palmer's (1996) Framework of Task Characteristics and Li's (1996) Test Point Validity Theory, this paper aims to analyze the test point quality and content validity of grammar gap-filling tests of the 2014–2019 National Matriculation English Test paper II (NMET II), in particular, concerning the content validity of tests in terms of the characteristics of input and expected responses, and test point validity. The results show that the grammar gap-filling tests in the NMET II are all in line with the language knowledge part required in the New Curriculum Standard and the Testing Syllabus and have a relatively high content validity. The test points are evenly distributed and reasonably designed, which can test students' grammar ability under the guidance of discourse with well designed length and reasonable test point validity. However, there are still some problems to be improved, such as the lack of 'man and nature' and practical essay writing in terms of topic and genre. The readability fluctuated greatly. The expected response type is restricted to limited production. Furthermore, test points designed with discourse level and meaning factor are less examined, and lacking of test points used to measure a higher grammatical ability.

Key Words: content validity, grammar gap-filling tests, NMET II

1. Introduction

As a compulsory subject of the college entrance examination, English has always been the focus of the reform. With the revision and adjustment of the New Senior English Curriculum Standards 2017 (NSECS 2017) for senior high school, the design of grammar test questions in English curriculum has also changed. In 2014, a grammar gap-filling test appeared in the National Matriculation English Test (NMET) for the first time, replacing the grammar multiple-choice test. Grammar gap-filling test is mainly used to measure students' comprehensive ability to use language knowledge in context. The changes of testing item in the college entrance examination of English not only affect students' scores, but also affect the cultivation of their language use ability. Therefore, it is particularly important to make an in-depth analysis of grammar gap-filling tests. As one of the criteria to measure the quality of a language test, content validity directly reflects whether the test can effectively measure what it should measure (Henning, 1987). Therefore, content validity is very important, because it can affect the washback effect of the test and the consequence of language teaching and learning (Hughes, 2003; Zhenjie Liu & Minghao Jin, 2018). As a large-scale standardized test, the NMET has a direct impact on the achievement of merit-based enrollment.

* This research was supported by the research funds of "The 13th Five Year-Plan" Key Research Project of Social Sciences of Education Department of Jilin Province ([2016], No. 259). Ming-Hao Jin received the fund in 2016.

** The first author is Ming-Hao Jin, and the corresponding author is Mei-Xian Zheng

2. Theoretical Background

As the college entrance examination of English subject, the NMET II is a large-scale, high stake, nationwide talents selection test. In order to measure students' grammatical ability in a real context, the grammar gap-filling tests took place the original grammar multiple-choice question test from the 2014 NMET II. The score of grammar gap-filling test is 15, accounting for 10% of total score of the NMET II.

The scholars who have studied the theoretical framework of grammar testing believe that the grammar gap-filling test is not limited to the examination of grammar knowledge alone, but the examination of the students' grammar ability in a certain language environment. For example, Halliday & Hasan (1985), Larsen-Freeman (1991), Alderson (2000), and Purpura (2004) all emphasized that grammatical knowledge should be expanded in discourse, that is, grammatical knowledge and grammatical ability should not be examined in sentences in isolation, but should be examined under the support of discourse. Compared with the grammar multiple-choice test, the grammar gap-filling test combines vocabulary, sentence and meaning, which can effectively test the basic vocabulary and grammar knowledge of the examinees, thereby testing the examinees' actual language ability.

The validity study of English grammar gap-filling test items in Chinese college entrance examination can be divided into diachronic research and synchronic research, based on the research results from C. Zhang (2015), H. Zhang (2016), Liu (2016), Guo (2017), Wang & Yang (2018), and Liu (2018), etc.

By sorting out the above literature, this paper finds that foreign studies on the theoretical aspects of grammar knowledge are relatively mature, but there are few studies on the content validity of grammar gap-filling tests. However, the number of domestic studies on grammar gap-filling tests has been increasing year by year since the change from grammar multiple-choice tests to grammar gap-filling tests in 2014.

At present, there are abundant researches on the validity of the NMET, but there is still a lack of research on analyzing the quality of grammar gap-filling tests in the NMET II, so as to put forward suggestions for grammar instruction in senior high schools. In view of this situation, this paper aims to compare the grammar gap-filling tests of the 2014–2019 NMET II with the requirements of the New Curriculum Standards and the Testing Syllabus, in particular, concerning the quality of test paper and content validity of the tests in terms of test description characteristics, the characteristics of input and expected responses, and the test point validity, based on Bachman and Palmer's (1996) Framework of Task Characteristics and Li's (1996) Test Point Validity Theory.

3. Research Design

3.1. Research Questions

This study aims to investigate the content validity of grammar gap-filling tests from the 2014–2019 NMET II, and the research questions are as follows:

- (1) What are the test description characteristics (including test requirements and structures)?
- (2) To what extent do the characteristics of input and expected response of the 2014–2019 NMET II grammar gap-filling tests match the requirements of the New Curriculum Standard and the Testing Syllabus?
- (3) To what extent does the test point validity of the 2014–2019 NMET II grammar gap-filling tests match the requirements of the New Curriculum Standard and the Testing Syllabus?

3.2. Research Objectives

This paper focuses on the research object of the grammar gap-filling tests of the NEMT II from 2014 to 2019, and studies its characteristics of input and expected response as well as test point validity, so as to measure the content validity of the gap-filling tests of the NMET II in the past six years.

3.3. Research Framework

Based on Bachman and Palmer's (1996) Framework of Task Characteristics and Li's (1996) Test Point Validity Theory, this paper examines the matching degree between six grammar gap-filling tests in the 2014–2019 NMET II and the requirements of language knowledge in the NSECS (2003 & 2017) and the Testing Syllabus. The research was conducted from three dimensions: test description characteristics (including test requirements and structure), characteristics of input and expected response (including genre, theme, length, readability, and expected response type), and test point validity (including test point distribution, test point level, and focus factors). The research framework is shown in Table 1.

Table 1. Research Framework

Test description characteristics	Test requirements	According to the Framework of task characteristics by Bachman(1996)
	Test structures	According to the Framework of task characteristics by Bachman(1996)
Characteristics of input and expected response	Genre	Narrative, expository, argumentative, practical writing
	Theme	Man and self, man and society, man and nature
	Length	Total number of words (excluding article titles, Chinese characters in brackets and Spaces)
	Readability	Take the Flesch reading ease table as the standard
	Expected response type	Selected, limited production, extended production
Test point validity	Test point distribution	The 24 grammar test points stipulated in the 2019 syllabus
	Test point level	Word level, phrase level, sentence level, discourse level
	Focus factors	Grammar, idiomatic collocation, meaning

4. Results and Discussion

4.1. Test Description Characteristics

According to Bachman and Palmer's (1996) test instructions in the test description: First, the test instructions of the NMET II are given in Chinese, and the test contents are given in English; Second, the NMET II in 2014 discarded the grammar multiple-choice items but used grammar gap-filling test items, which measures the ability of the examinees to use the knowledge of English grammar in discourse, that is, the ability of the examinees to use the grammar in a discourse context by reading a passage of about 200 words and filling in the blanks with appropriate words or the correct form of the words in brackets. The variation of the test instructions of grammar gap-filling tests from the 2014–2019 NMET II are shown in Table 2.

Table 2. Test Instructions of the 2014-2019 NMET II Grammar Gap-filling Tests

	Test instructions
2014	Read the material below and fill in the blanks with the appropriate content (no more than 3 words) or the correct form of the word in brackets.

2015	Read the material below and fill in the blanks with the appropriate content (1 word) or the correct form of the word in brackets.
2016	Read the material below and fill in the blanks with the appropriate content (1 word) or the correct form of the word in brackets.
2017	Read the passage below and fill in the blanks with an appropriate word or the correct form of the word in brackets.
2018	Read the passage below and fill in the blanks with an appropriate word or the correct form of the word in brackets.
2019	Read the passage below and fill in the blanks with an appropriate word or the correct form of the word in brackets.

According to the 2014–2019 NMET II grammar gap-filling test instructions shown in Table 2, the test instruction underwent two changes. The first was to change "(no more than 3 words)" to "(1 word)" in the instruction in 2015 and 2016, possibly because more explicit instructions were used to avoid the possibility of affecting students' scores. The second change was made from 2017 in two aspects and its reasons are as follows: 1) In 2017 NMET II, the word "material" was changed to "passage" in test instruction. Compared with the previous NSECS (2003), the NSECS (2017) added thematic context and discourse types, enabling students to learn by discourse under the guidance of the theme, instead of learning vocabulary and grammar in isolation, which is conducive to forming an integrated and related view of English subject learning. 2) Change the word "content" to "word" from 2017 test instruction. The NSECS (2017) states that the form of test paper should be standardized, the instructions should be clear and concise, and that the way of answering questions should be clarified. In order to better guide the examinees to understand the requirements of the test, the test instructions of grammar gap-filling tests are changed from the original "fill in the blanks with the appropriate content" to "fill in the blanks with an appropriate word" from 2017 test instruction of grammar gap-filling tests.

4.2. Characteristics of Input and Expected Response

4.2.1. Topic and Genre

According to the basic knowledge of language part required in the NSECS (2003) and the NSECS (2017) for students in senior high school, the comparison of topic and genre in grammar gap-filling tests from the 2014–2019 NMET II were made and shown in Table 3.

Table 3. Topic and Genre of the 2014- 2019 NMET II Grammar Gap-filling Tests

	Theme	Topic	Genre
2014	Man and self	A little boy returned the suitcase.	Narrative
2015	Man and society	American Indians built adobe houses to regulate the characteristics of the greenhouse.	Expository
2016	Man and self	How to relieve pressure at work?	Argumentative
2017	Man and society	The development of the first underground passenger railway into a subway.	Expository
2018	Man and society	China is adjusting the structure of its crops because of changes in people's diets.	Expository
2019	Man and self	90-year-old woman wins 'woman of the year' award for working every week.	Narrative

From Table 3, the 2014 topic of "A little boy returned the suitcase." as well as the 2016 topic of "How to alleviate the pressure of the work?" is closely related to daily life, the 2015 topic of "American Indians built adobe houses to regulate the characteristics of the greenhouse." shows the local customs and culture. The genres from 2014–2016 are narrative writing, expository writing, and argumentative writing respectively so that the genres are diverse and not repetitive, all in line with the requirements of the NSECS (2003).

According to the NSECS (2017), there are three thematic contexts and four genres of writing. They are topics of 'man and self', 'man and society', 'man and nature' and genres of narrative, expository, practical and argumentative writing.

The topics of the 2014–2019 NMET II grammar gap-filling tests involve the thematic contexts of 'man and self' and 'man and society' in the NSECS (2017), which contains excellent character, material and cultural heritage, healthy lifestyle, social progress, positive life attitude, etc., which can enable students to learn important contents of language, enrich their life experience and thinking mode, and influence their thinking development level as well. At the same time, through understanding of culture and historical development, students can pay attention to cultural history and improve cultural awareness. Release of the NSECS (2017), therefore, not only accepted the requirements of the NSECS (2003), but also divided the theme context into three types: man and self, man and society, and man and nature, and also divided the major thematic contexts into different theme groups, and then put forward more specific sub-themes, so as to further enrich and improve the curriculum standards.

However, it is found that the thematic context of 'man and nature' and the genre of practical writing have not been involved in the 2014–2019 NMET II grammar gap-filling tests. It is suggested to design grammar gap-filling tests with the topic related theme of 'man and nature' and the genre of practical writing to promote the comprehensive development of students' language knowledge, thinking quality and learning ability, and help them communicate and express themselves through reading different genres of context, so as to better meet the requirements of the NSECS (2017).

4.2.2. Length

Bachman (1990) pointed out that the length of the input and response, composed of language samples, may range from a single word, to a sentence, to an extended chapter. While the length of the language sample itself may not be the key factor affecting the examinees' performance, the longer the language sample, the greater its potential impact on other input and response characteristics. In this paper, the text length refers to six articles in the grammar gap-filling tests of the 2014–2019 NMET II. To make the collected data more objective and accurate, this study uses Microsoft Office for Mac 2016 to calculate the length of the article. The text length of grammar gap-filling tests of the 2014–2019 NMET II is shown in Table 4.

Table 4. Text Length of the 2014-2019 NMET II Grammar Gap-filling Tests

	2014	2015	2016	2017	2018	2019
Text length	196	163	184	197	196	196
Difference from the reference length	-4	-37	-16	-3	-4	-4
Percentage difference	2.0%	22.7%	8.7%	1.5%	2.0%	2.0%

According to the description of the 2019 Testing Syllabus, a 200-word essay with 10 blank spaces requires examinees to fill in the blanks with appropriate words, or the correct form of the words in brackets, to test their ability to use grammar based on reading comprehension. According to the analysis of the 2014–2019 NMET II, the minimum length of the text is 163 words in 2015, and the maximum length is 197 words in 2017. From 2014 to

2019, the average text length of grammar gap-filling tests of the NMET II is 189 words, among which the text length in 2015 and 2016 is significantly less than 200 words, and the text length in 2014, 2017, 2018 and 2019 is designed with nearly 200 words, which meets the requirement the Testing Syllabus.

4.2.3. Readability

In this paper, Flesch Reading Ease and Flesch-Kincaid Grade Level, which are the most commonly used and with high accuracy, are adopted to detect and analyze the readability of the articles on grammar gap-filling tests of the NMET II from 2014 to 2019. Readability is divided into seven levels, corresponding to the group with seven different levels of reading ability, among which the readability of articles of standard difficulty should be 60–70 (Alderson, 2000). The score mapping reference table of Flesch Reading Ease is shown in Table 5.

Table 5. Reference Table of Flesch Reading Ease

Readability	Difficulty level	Grade (native speaker)
0–30	Very difficult	College
30–50	Difficult	High school
50–60	Fairly difficult	Partial high school
60–70	Standard	Seven or eight
70–80	Fairly easy	Six
80–90	Easy	Five
90–100	Very easy	Four

In order to be more accurate, this paper uses Microsoft Word 2010 to calculate the readability of six articles of grammar gap-filling tests. Through comparative analysis of the calculation formula of readability and the software display of readability, the readability values of the six articles in the grammar gap-filling tests of the 2014–2019 NMET II are obtained, as shown in Table 6.

Table 6. Readability of the 2014-2019 NMET II Grammar Gap-filling Tests

Year	Readability	Difficulty level
2014	82.9	Easy
2015	62	Standard
2016	68.9	Standard
2017	59.8	Fairy Difficult
2018	45.4	Difficult
2019	69.2	Standard
Average	64.7	Standard

As shown in Table 6, the readability of the NMET II grammar gap-filling tests from 2014 to 2019 ranges from 45.4 to 82.9, with an average readability of 64.7, which assesses as ‘standard’ according to Table 5. In 2014, the

readability of grammar gap-filling test exceeded 80, indicating that the difficulty level is 'easy'. The reason is that grammar gap-filling test was used for the first time in 2014 to replace the original grammar-multiple choice tests. It is necessary to consider the degree of adaptation of examinees to the new type of grammar test, so the difficulty was reduced. In addition, the readability fluctuated greatly from 2017 to 2019. Therefore, according to the requirements of standard difficulty level articles of the reference table of Flesch Reading Ease, it is suggested to control it around standard level 60-70.

4.2.4. Expected Response Type

Bachman and Palmer (1996) pointed out that there are three types of expected responses, namely selective production, limited production and extended production. The selective production asks the examinees to choose the correct or wrong answer among several choices. Limited production asks the examinees to write down a word, phrase or sentence to complete a sentence. An extended production is a response that consists of several sentences and asks the candidate to write down several sentences. Among the three types of responses, expanded production has the highest requirement on the examinees' language ability, while selective production has the lowest requirement.

Thus, it can be seen that the grammar gap-filling tests of the NMET II from 2014 to 2019 belong to the limited production response type, that is to say, word gap-filling and sentence completion is the only test form. The NMET II tested in 2014 is used to replace multiple-choice question type grammar test with grammar gap-filling, requiring examinees to fill in the correct tense and voice of the given word or the appropriate form of the word according to context. Although multiple-choice question type grammar tests used in and before 2013 NMET II can ensure the objectivity and reliability of the test, and save time and manpower, they can't directly measure examinees' real language ability because of the lack of interaction and communication. Furthermore, it is inevitable that examinees may guess the answer.

Therefore, after the reform and adjustment into grammar gap-filling test, examinees are required to analyze the different grammar test points in the text according to the text clues. To a large extent, it meets the requirements of the NSECS (2017) on grammar knowledge, that is, grammar knowledge is the unity of 'form, meaning and use', and meanwhile it guides the examinees to understand and use of grammar knowledge in an appropriate context so as to further improve their awareness of English grammar.

After the reform of the test type of grammar knowledge in 2014, compared with the previous scattered knowledge points, the new test type has a specific entry point for the examination of grammar knowledge and a clear positioning for the examination review. From the multiple-choice questions before 2014 to the grammar gap-filling tests from 2014, that is, from the selective production response to the limited production response, although the college entrance examination of grammar knowledge has made a qualitative progress, it should also appropriately increase the type of expanded production response of the test in the future, which can better reflect the purpose of the college entrance examination to select outstanding talents.

4.3. Test Point Validity

4.3.1. Test Point Distribution

The 2019 Testing Syllabus provides 24 grammar test points, including nouns, pronouns, numerals, prepositions, conjunctions, adjectives, adverbs, articles, verbs, tenses, voice, the predicate verb, word formation, sentence types and sentence components, basic sentence patterns of simple sentences, subject-predicate agreement, complex sentence, indirect speech, ellipsis sentence, inverted sentence, emphatic sentence and subjunctive mood. This paper analyzed the distribution of grammar gap-filling tests of the 2014-2019 NMET II in terms of those 24 grammar test points mentioned above, as shown in Table 7.

Table 7. Test Point Distribution of the 2014-2019 NMET II Grammar Gap-filling Tests

Classification	Test point	2014	2015	2016	2017	2018	2019
Vocabulary	Noun		1(63)	2(62, 66)	2(61, 69)	1(66)	
	Pronoun	1(69)			1(67)		
	Preposition	1(64)	1(67)	1(64)	1(62)	1(65)	
	Conjunction	1(62)					1(67)
	Adjective	1(63)	1(69)	1(61)	1(70)	1(67)	1(70)
	Adverbs	1(70)	1(65)	2(65, 67)	1(66)	1(63)	1(63)
	Articles		1(62)	1(68)	1(64)	1(62)	1(69)
Grammar	Tense	2(65, 68)	1(68)	1(70)	2(65, 68)	2(61, 68)	2(64, 66)
	Voice				1(65)	1(61)	
	Non-predicate verb	3(61, 66, 67)	3(61, 64, 65)	1(69)	1(63)	2(64, 70)	3(61, 65, 68)
	Subordinate relatives		1(70)			1(69)	1(62)
	Subject-predicate agreement			1(63)			1(66)

Note: 2 (65, 68) = '2' means the number of test points, '65, 68' means the sequence of questions, '65' means more than one test point.

As can be seen from Table 7, the ten spaces in the grammar gap-filling tests are divided into two types: One is the blank space with prompt words, including the agreement of verb tense, voice and subject-predicate verbs, nouns, and the part of speech conversion of adjective and adverb. The other is the blank space without prompt words. The answer can only be one word, including pronouns, articles, prepositions, conjunctions and relative words of clauses. The grammar examination points of grammar gap-filling tests of the 2014-2019 NMET II are distributed widely and evenly, mainly including nouns, pronouns, prepositions, conjunctions, adjectives, adverbs, articles, tenses, voice, non-predicate verbs, subordinate relative words and subject-verb agreement, which are all within the scope of the grammar test points specified in the Testing Syllabus. It is found that the 2014-2019 NMET II grammar gap-filling tests put their focus on the application of basic grammar and vocabulary knowledge, involving 7-9 related test points in average, which meet the requirements of the Testing Syllabus.

However, only 13 of the 24 grammar points are covered in the past six years' grammar gap-filling tests. There is no examination on 11 other test points, especially indirect speech, ellipsis sentence, inverted sentence, emphatic sentence, subjunctive mood, etc., which are used to measure students' higher grammatical ability. It may lead to teachers and students pay less attention to these test points in classroom teaching and learning, and cause negative washback effect on developing students' comprehensive language use ability.

4.3.2. Test Point Level

According to the classification of test points by Xiaoju Li (1997), this paper discusses the test points of grammar gap-filling tests from four aspects: word level, phrase level, sentence level and discourse level. Word level is the lowest level and does not require contextual support. The phrase level refers to the words that usually appear in collocation form, selected by the information before and after the blank space. Sentence level refers to the words that are determined in a sentence. In the discourse level, a sentence does not determine the word, but the content of the whole text. Therefore, the discourse level is the highest among the four levels. The higher the level of the test point, the stronger the language ability is needed. Therefore, if we want to measure students' real language ability, it is needed to improve the test point level of the test to improve the efficiency. This paper analyzed the test level of 60 grammar gap-filling tests of the 2014-2019 NMET II, as shown in Table 8.

Table 8. Test Point Level of the 2014-2019 NMET II Grammar Gap-filling Tests

Test level	2014	2015	2016	2017	2018	2019
Word level	0	0	0	0	0	0
Phrase level	5(61, 64, 66, 67, 70)	5(62, 63, 64, 67, 69)	7(61, 62, 64, 65, 67, 68, 69)	6(61, 62, 64, 66, 69, 70)	4(62, 65, 66, 67)	4(61, 65, 67, 69)
Sentence level	3(62, 63, 68)	5(61, 65, 66, 68, 70)	2(66, 70)	4(63, 65, 67, 68)	6(61, 63, 64, 68, 69, 70)	6(62, 63, 64, 66, 68, 70)
Discourse level	2(65, 69)	0	1(63)	0	0	0

Table 8 shows that the 2014–2019 NMET II has not involved the test points of word level, while phrase level and sentence level test points are the main part, and only three involve the discourse level. It is worth noting that the phrase level test points are more than sentence level ones in the grammar gap-filling tests from 2014 to 2017, while the sentence level test points are more than word level ones in the grammar gap-filling tests from 2018 to 2019. The reason is that, the new version of English Curriculum Standards released in 2017, the NSECS (2017), which emphasizes that students should learn English from the thematic context instead of learning vocabulary and grammar in isolation.

Since are only three discourse level test points in the 2014–2019 NMET II, two in 2014, and one in 2016, in order to improve the validity and discrimination of the test, it is suggested to design certain proportion of test points designed with discourse level.

4.3.3. Focus Factor

The focus factor is closely related to the test point level. The focus factor refers to the different factors of collocation, meaning and rhetoric that should be examined on each test point level (Xiaojuli, 1997). The focus factors can be divided into meaning factors, idiomatic collocation factors and grammatical factors. The corresponding focus factor of each test point level is also different. Low-level questions measure grammatical factors, while high-level questions measure meaning factors. If a test point is used to measure meaning factor, it can also measure grammatical factor as well, but not the other way around. Table 9 shows the distribution of focus factors of grammar gap-filling tests of the 2014–2019 NMET II.

Table 9. Focus Factors of the 2014-2019 NMET II Grammar Gap-filling Tests

Focus factor	2014	2015	2016	2017	2018	2019	Percent
Meaning factor	69			67			3.3%
Idiomatic collocation factor	61, 64, 66, 67	67	64, 65, 68, 69	70	65	67	20%
Grammatical factor	62, 63, 65, 68, 70	61, 62, 63, 64, 65, 66, 68, 69, 70	61, 62, 63, 66, 67, 70	61, 62, 63, 64, 65, 66, 68, 69	61, 62, 63, 64, 66, 67, 68, 69, 70	61, 62, 63, 64, 65, 66, 68, 69, 70	76.7%
Total	60 Test items						100%

As can be seen from Table 9, there are a total of 60 grammar gap-filling tests in the NME II from 2014 to 2019. Grammatical factors are the most frequently examined, accounting for 76.7% of the total number of questions. The second is the idiomatic collocation factor, accounting for 20%. Only one meaning factor is involved in a test point in 2014 and 2017 grammar gap-filling tests respectively, accounting for 3.3%. Thus, it can be seen that the distribution of grammatical factors is relatively stable in terms of the distribution of focus factors in the 2014–2019 NMET II grammar gap-filling tests, while the proportion of the idiomatic collocation and meaning factors is unbalanced.

5. Conclusion

This paper conducted a diachronic study on the content validity of the 2014–2019 NMET II grammar gap-filling tests. The results show that the grammar gap-filling tests of the 2014–2019 NMET II meet the requirements of grammar knowledge specified in the Testing Syllabus and the English Curriculum Standards. Thus, the grammar gap-filling tests have a high content validity.

It is found that the grammar gap-filling tests can not only measure students' basic grammar and vocabulary knowledge, but also students' ability of using grammar in thematic context. The test points are widely and evenly distributed with a reasonable length and the average readability is 64.7, which meets the standard. However, there are some deficiencies should be considered in designing the grammar gap-filling test, such as lacking the topic of 'man and nature', genre of practical writing, 11 other test points, especially those which can be used to measure higher grammatical ability, and inconsistent of readability, designing expected answers with limited production response only, less examined test points in terms of discourse level and meaning factors, etc.

References

- Alderson, J. C. (2000). *Assessing reading*. Cambridge: Cambridge University Press.
- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice*. Oxford: Oxford University Press.
- Examination Center of the Ministry of Education. (2019). *The national college entrance examination test syllabus of 2019*. Beijing: Higher Educational Press.
- Guo, B. (2017). Contrastive analysis and enlightenment of English grammar cloze and essay correction in college entrance examination. *Educational Measurement and Evaluation*, 5, 18–23.
- Halliday, M. A .K., & Hasan, R. (1985) *Language: Context and text*. Burwood: Deaken University.
- Hughes, A. (2003). *Testing for language teachers* (2nd ed.). Cambridge: Cambridge University Press.
- Larsen-Freeman, D. (1991). *Teaching grammar*. In M. Celce-Murcia (Ed.). *Teaching English as a second or foreign language* (pp. 251–266). Boston MA: Heinle & Heinle Publishers.
- Liu, H. (2016). A construct validation study of grammar gap-filling task in NMET Guangdong. Unpublished MA thesis. Guangdong: Guangdong University of Foreign Studies.
- Liu, S. (2018). A study in the task characteristics of NMET grammar gap-filling task. *Journal of Nanchang Education College*, 2, 20–23.
- Li, X. (1997). *Science and art of language testing*. Changsha: Hunan Education Press.
- Liu, Z., & Jin, M. (2018). A study of content validity of reading comprehension tests in NMET II. *The Linguistic Association of Korea Journal*, 26(4), 103–124.
- Ministry of Education of the People's Republic of China. (2003). *National English Curriculum Standards for general High School* (Experimental edition). Beijing: People's Education Press.
- Purpura, J. (2004). *Assessing grammar*. Cambridge: Cambridge University Press.

- Wang, X., & Yang, Z. (2018). A study on content validity of the application of examination items about English language knowledge from National Matriculation English Test in 2017. *Educational Measurement and Evaluation*, 4, 25-33.
- Zhang, C. (2015). A Rasch-based construct validation of GMET grammar gap-filling task. *Modern Foreign Language*, 6, 126-127.
- Zhang, H. (2016). Test method effects in grammar tests: The comparative study between grammar gap-filling tests and multiple-choice tests. *Foreign Language Testing and Teaching*, 2, 32-38.

Ming-Hao Jin

Associate Professor

Department of English, Yanbian University

977 # Gongyuan Road, Yanji City, Jilin Province, China

Phone: +86-433-273-2742

E-mail: jinminghao@ybu.edu.cn

Mei-Xian Zheng

Graduate Student

Department of English, Yanbian University

977 # Gongyuan Road, Yanji City, Jilin Province, China

Phone: +86-139-4470-4768

E-mail: mihyon0201@live.com

Received on July 29, 2020

Revised version received on October 15, 2020

Accepted on October 16, 2020