A Study on the Content Validity of English Writing Tests of GDNMET in the Circumstance of Ten-Year New Curriculum Reform

XU Jie[a],*

[a]Ph.D., School of Training, Lingnan Normal University, Zhanjiang, China.
*Corresponding author.

Abstract

Content validity is primary and essential to test. Writing covers 26.7% in English test in GDNMET and 20.8% in NMET. It will be meaningful to see whether English writing tests in the past 10 years have high validity or not by answering these questions: Is the writing content related to ECSCSMS and testing syllabus? Are the writing items representative? Is the writing content suitable for candidates of GDNMET. As a result, according to multidimensional analysis, English writing tests of GDNMET from 2007 to 2016 have high content validity. What's more, there are some implications on the coming testing reform of GDNMET and English teaching in senior high schools.

Key words: Content validity; English writing test; New curriculum reform; English teaching

INTRODUCTION

No language testing is perfect. No test is believed as perfectly standardized as its objectives. In recent years, much has been written on language testing (e.g. Alderson et al., 1991, 1995; Bachman, 1990; Bachman & Palmer, 1996; Heaton, 2000; Henning, 2001), especially on the issue of validity, but some problems in testing implementation still exist. As a province-wide test that boasts an annual attending population of about 600,000 every year, GDNMET calls for the attention of not only teachers and students, but also people from all walks of life, having a great social impact. English, as one of the key subject, will inevitably exert great impact on its teaching and learning in high schools. Like any test, it should be reliable and valid for the purpose for which it is intended. Besides, the greater a test's social impact is, the higher its need for validity. Considering the substantial amount of washback this test generates, the need for validation studies seems all too evident.

For promoting Chinese high school English teaching reform and development, Ministry of Education promulgated the English Curricular Standard for Common Senior Middle School (ECSCSMS) in 2003. Since September, 2004, the new curriculum reform has been carried out firstly in Shandong, Ningxia, Guangdong and Hainan these four experimental areas. And it was the first time that senior high school students of Guangdong experienced new curricula and had the National Matriculation English Test (NMET) of Guangdong Province (GDNMET) in 2007. Ever since then, ten years have passed. And according to 2017 New Testing Reform Plan, English will be tested several times a year and the full marks will stay the same, that is still scored 150 points.[1]

Writing test is a kind of comprehensive test. It not only tests the language knowledge of vocabulary, grammar and usage, but also the ability of reading, summarizing, conceiving, organizing and expressing,

---

[1] 2017 New Testing Reform Plan. [DB/OL]http://www.360doc.com/content/14/0911/17/8224881_408716343.html

etc. It can truly reflect the actual language level of students and also has positive backwash on English teaching. Very little research has been devoted to the content validity of the writing test of a provincial NMET in the circumstance of new curriculum reform. For this reason, I have carried out the study on the writing part in GDNMET from 2007 to 2014, to find out whether the eight writing tests have high content validity or not, hoping that my research findings might shed some light on the next testing reform in 2017 and provide a little inspiration and experience for English teaching in senior high schools.

1. WHAT IS VALIDITY AND CONTENT VALIDITY?

1.1 Validity

Validity is considered to be one of the most important notions in testing. Bachman (1990, p.24) pointed out the validity is "essential to the interpretation and use of measures of language ability." It is the primary and necessary quality to be considered in developing and using tests. What's more, Heaton (2000, p.164) considered test validity as one of "chief criterion for evaluation any test, whatever the theoretical assumptions underlying the test." Henning (2001, p.89) also believed: Validity in general refers to the appropriateness of a given test or any to its component parts as a measure of what it is purposed to measure. From their points of view, it is not difficult to see how important validity is for language testing.

1.2 Content Validity

Generally speaking, there are several kinds of validity: Content validity, face validity, construct validity, concurrent validity and predictive validity, among which content validity is considered to be "the most difficult to understand and also the most important." (Alderson, Clapham, & Wall, 2000, p.172)

Content validity is concerned with whether or not the content of the test is sufficiently representative and comprehensive for the test to be a valid measure of what it is supposed to measure (Henning, 2001, p.94). In order

to evaluate content validity of a language test, a careful analysis of language being tested is needed. Wood (2001, p.148) states "content validity is a function of success with which the board translates what the syllabus says into an operational examination. If the content is not validly selected, then score interpretations are liable to distortion."

To make a conclusion, for the content validity, what is going to do is to check whether the content measured in the test reflects enough the content required to be measured in the testing syllabus.

2. OBJECTIVES OF STUDY

Based on Liu and Han (2000, p.208), three factors should be considered in order to value content validity of a test: a) whether the test content is related to test purposes; b) whether the test items are representative; c) whether the test content is suitable for test-takers.

So, as for the content validity of writing, these objectives of this study on the writing tests are to find out: a) whether the writing content is related to ECSCSMS (English Curricular Standard for Common Senior Middle School) and testing syllabus; b) whether the writing items are representative; c) whether the writing content is suitable for candidates of GDNMET. Therefore it is possible to find out whether the writing test from 2007 to 2015 in GDNMET and 2016 in NMET has high content validity or not.

That is to say, in order to tell the content validity correctly, ECSCSMS, testing syllabus, test papers, their possible versions and scoring criteria, even the official annual report on GDNMET are necessary to be analyzed in detail.

3. DISCUSSION AND FINDINGS

There are two writing items in the test from in GDNMET, one is basic writing, which means to control writing; the other is task-based writing, which means divergent writing. These two different testing items have different requirements and purposes. Whereas, there is only one writing item in NMET.

Table 1
Overview of Writing Test in GDNMET

Item	Time	Value	Requirements	Testing purpose
Basic writing	10 mins	15 points (10%)	Use only five sentences to write a passage according to the given information and requirements given. Students can not omit any information. No words requirements.	To test the ability of using the right and suitable sentences to express given information.
Task-based writing	25 mins	25 points (16.7%)	Students are required to read an English passage first and then sum up the main idea in about 30 words. Then according to the requirements given in Chinese, students should write a composition which is related to the topic of reading by using 120 words.	To test comprehensive language use ability, especially the ability of analyzing, summarizing and expressing.

Table 2
Overview of Writing Test in NMET

Item	Time	Value	Requirements	Testing purpose
Writing	30 mins	25 points (16.7%)	Students are required to use appropriate vocabulary, grammar and sentence structure to convey information in a clear and coherent way by using 100 words.	To test comprehensive language use ability, such as the ability of analyzing and expressing, the ability of using appropriate expressions to convey information.

3.1 Is Writing Content Related to ECSCSMS and Testing Syllabus?

Test writers develop the test, following the testing syllabus which is considered to be essential in establishing test's validity. The development of the testing syllabus is "a central and crucial part of the test construction and evaluation process." (Alderson, Clapham, & Wall, 2000, p.9). Therefore, above of all，the content validity should be proceeded following the testing syllabus, which is usually affected by the curricular standard. If so, this test will have high content validity. As Table 1 shows, different ways of writing definitely improve testing scale, and it is good to test more required skills according to ECSCSMS, comparing with one item.

Based on the ECSCSMS, the overall objectives of English writing are as follows (Table 3). The guideline of GDNMET is that it should follow the rules and requirements of the recruitment exam, implement the conception of new curriculum, reflect the overall requirements of the new curricular standard, examine the candidates' preliminary scientific literacy and humanities, pay attention to examining both ability and quality, both time and practice to promote the implementation of quality education. Based on the guideline of the new curricular standard, Testing Syllabus issued testing requirements of GDNMET. Candidates are required to have the following writing skills (see Table 4).

Table 3
Requirements of ECSCSMS in Writing and Testing Syllabus

Requirements of ECSCSMS in writing[2]	Requirements of testing syllabus[3]
1.Students are able to rewrite or write the summary according to the given reading material; 2.Students are able to write a passage or report according to the given words or charts; 3.Students are able to write a passage with coherent content and complete structure, stating events and expressing opinions and attitudes ; 4.Students are able to pay attention to the normalization of the article genre and the smoothness of sentences.	Candidates are required to write according to the given information. 1.Be able to convey and express the information clearly and coherently; 2.To use the language knowledge effectively.

Table 4
Analysis of Genre and Topic of Testing Items in GDNMET

Year	Genre of basic writing	Topic of basic writing	Genre of task-based writing	Topic of task-based writing
2007	Letter	The survey about the idol	Narrations and comments	A father wrote a letter to his son when the son was born.
2008	Object introduction	Introduce shooting	Narrations and comments	How Mike adjusts himself to his college life.
2009	Activity report	The investigation about the students' short-sightedness	Argumentation	Koalas are unhappy and complaining about tourists taking photos.
2010	Activity report	The activities to prevent smoking in public	Argumentation	Different views about the phenomenon that give the children money to encourage them to study.
2011	Book introduction	Introduce a book called *Battle Hymn of the Tiger Mother*	Narrative writing	Sam was sent to a boarding school, in which he felt homesick, so he could not focus his attention on his lessons.
2012	Character introduction	Introduce a legendary Australian	Narrations and comments	A student thanked a retired teacher for helping her and giving a great influence on her life.
2013	Pinion expression	Choose the volunteers who will be sent to the Mars	Narration and comments	The ordinary worker and jobs, such as cleaning the street, should be respected.
2014	Activity report	Look for time twins	Narration and comments	Miss McCarty's donation which was her life savings was for the students who really needed financial help.
2015	Science report	Introduce a new technology, DNA test	Narrations and comments	Class rules set up by Sally helped her gain respect from her students.

[2] The Ministry of Education. (2003). *English curricular Standard for common senior middle school.*
[3] The Ministry of Education. (2010). *Testing syllabus for senior high school students.* Retrieved from http://club.topsage.com / thread-497623-1-1.html

As Table 4 states, the genre of both basic writing and task-based writing is diverse, including narrative, practical writing，exposition and argumentation, which accounts for a larger and larger proportion, provides students with fully show individual character and talent platform, fully embodies the direction of the reform of college entrance examination. Writing tasks in NMET in 2016, as indicated in Table 5, mainly focus on letters. No matter which kind of proposition, writing task is provided clearly, some are given in Chinese, some are given graphically combined with tips, some is in a table. Students are possible to write in different expressing way according to the clear information.

Table 5
Analysis of Genre and Topic of Testing Item in NMET

Year	Genre of basic writing	Topic of writing
2016(I)	Letter	Write a letter to the foreign teacher to ask him for help with the part-time job application form and resume

The topics are closely linked to examinee's life, such as the experience of missing home, showing grateful feeling to the father, class rules, writing a letter to ask for help. Besides, some topics are very concerned by students and full of period flavor. For example: the donation of *Tiger Mother Battle Song* and Miss McCarty. These topics conform to student's cognition level, enables the majority students have the speech to say when they are writing and shows their writing ability fully.

3.2 Are the Writing Items Representative?

To decide whether the writing items representative, it is necessary to see whether the testing way has its characteristics and whether it is able to measure what it is supposed to measure.

From Table 1, it is easy to see the two writing items pay attention to test candidates' language output performance. Depending on the requirements of basic writing and task-based writing, the testing purpose which has close relationship with testing syllabus will be easy to realize. It can measure different writing abilities including summarizing, conveying information and expressing opinions natively, deciding candidates' actual language use ability.

Furthermore, in order to evaluate students' exact level well, it's necessary to have a suitable and scientific scoring criteria (Tables 6 and 7). Basic Writing is aimed to test the student's ability of using the right and standard language to express given information, including the rationality of using words and expressions, the complexity of the sentence structure, the correctness of the use of grammar, the integrity of the information, the consistency between the sentences. Based on the requirements of basic writing, some relevant information is provided for students. In the composition, given information points can't be omitted, but for language-expressing way, the order of information organization, what information to make up, the candidates has to think of space to show their real writing ability, which conforms to the requirements of ECSCSMS. Students have to use only five sentences without words limit to write a passage. Scoring method is the analysis, the range of 15 points is divided into five levels, each level's gradient is 3 points. Teachers can easily have a command of the scoring method.

Task-based writing combines reading and writing together. It is aimed to convey the conception of lifestyle context and the real task. It tests student's ability of reading as well as writing. The scoring criteria ask for accurate summary, standard expression, suitable content and the coherent discourse. As a subjective question, it well reflects the conception of new curriculum reform and requirements of "cultivating the ability of comprehensive language use" in ECSCSMS. The total points of task-based writing contain the scores of summarizing and writing these two parts. Summary part covers 5 points, and the body part 20 points. Scoring method is the hierarchical analysis, and each criterion covers 5 points.

Writing task in NMET is designed to check students' writing ability and evaluate their ability to use relevant knowledge to effectively convey information and complete communicative tasks. The scoring criteria cover the completeness of the information, the coherence of sentences and the correctness and variety of sentence structure.

Table 6
Scoring Criteria and Requirements of Writing Items in GDNMET

Item	Basic writing	Task-based writing
Scoring criteria	1. The completeness of the information 2. The coherence of sentences 3. The correctness and variety of sentence structure	1. The accuracy of summary 2. The normalization of language use 3. The appropriateness of content 4. The coherence of sentences

Table 7
Scoring Criteria and Requirements of Writing Test in NMET

Item	Basic writing
Scoring criteria	1. The completeness of the information 2. The coherence of sentences 3. The correctness and variety of sentence structure

Even though the writing items have strict requirements and scoring criteria, as a new testing type, these two different testing ways make the test have strong and close relationship with ECSCSMS and testing syllabus. Controlled writing as well as divergent writing is good to test examinees' spelling, sentence coherence, summary and awareness of discourse, letting the evaluation more comprehensive, multidimensional and representative.

3.3 Is the Writing Content Suitable for Candidates of GDNMET?

Whether writing content is suitable for candidates or not relies on the difficulty of the two items. If the writing content is over difficult or easy, it is impossible to get the discrimination between candidates.

Task-based writing is closely related to the reading material. Reading materials to the examinee to inspire, but is not the purpose directly to say it. Writing topics come from reading materials, but not simply repeat the content of the reading material, nor indefinitely extended topic from the material. So it is necessary to see whether the reading material provided influences the comprehending, which directly affects the summary and of body writing.

Table 8
Analysis of Reading Material of Task-Based Writing

Year	New words	Syllabus	Total words	Sentence	Average	Readable degree	Difficulty
2007	2	277	240	18	13.3	96.05	Very easy
2008	1	248	187	12	15.6	78.49	Ok
2009	2	244	169	8	21.1	63.6	Not difficult
2010	1	241	188	13	14.5	83.83	Easy
2011	1	235	192	14	13.7	89.72	Easy
2012	1	266	193	12	16.1	73.75	Ok
2013	1	268	190	10	19	68.26	Not difficult
2014	1	305	208	13	16	66.24	Not difficult
2015	0	298	210	14	15	71.61	OK

In the reading material, new word quantity are controlled below 3% in line with the requirements of the testing syllabus. All the reading materials are eligible as for this condition. Readable degree formula "RE=206.835−(0.846*NSYLL) −(1.015*W/S)" (Alderson, 2000) is a good way to decide whether a reading material is easy or difficult for candidates.

NSYLL refers to the number of syllabus in one hundred words, while W/S is for the average sentence length of the reading material. The lower score is, the more difficult the reading will be. From Table 8, the decisive factor, readable degree is telling that the reading materials are very suitable for candidates to understand.

Table 9
Analysis of Given Possible Versions of Basic Writing

Year	2007	2008	2009	2010	2011	2012	2013	2014	2015
Length	88	77	94	80	105	124	152	92	85
Coherence words	4	1	3	1	2	3	3	1	1
Sentence patterns[4]	L*3,C*2,M*2, A3*2,A4	F*2,L	A4,A6, B1,B2, C,D,F,M	B1*2, I,L	A3,A6, B1,B2, F*2,G, H,M,L	B2*2,I, G,A5,B1	G,B1*, L,M,F, A2*2, I,A1	F,B1, A6,A2 ,L	L, C*3
Compound sentence	4	3	4	2	3	3	4	5	1

---

[4] PS: A1. The adverbial clause of condition; A2. The adverbial clause of time; A3. The adverbial clause of concession; A4. The adverbial clause of cause; A5. The adverbial clause of result; A6. The adverbial clause of purpose; B1. The restrictive attributive clause; B2. The non-restrictive attributive clause; C. Object clause; D. Predicative clause; E. Subject clause; F. Appositive clause; G. It as a formal subject; H. Emphatic sentence; I. Inverted sentence; J. The subjunctive mood; K. With + Compound construction; L. Compound sentence; M. Elliptical sentence; N. Exclamatory sentence

Table 10
Analysis of Given Possible Versions of Task-Based Writing

Year	2007	2008	2009	2010	2011	2012	2013	2014	2015
Summary	39	39	32	29	34	37	41	34	45
Length	209	197	166	224	163	245	238	189	183
Coherence word	1	7	4	5	7	7	6	4	4
Sentence patterns	A2*3,C,H,M, N,B2,A5	B2*2,E,I,C,B2, G*2,D,L	F*2,B1, A4,LG*3,C,J, A1,F,A6,I*3	E,L*5,A6*2, C,G,A2,K,C, B1,F,I	F,B2,A5,L*2, A2*2, A4*3, A3*1,G,C*2	B2*4,B1,A2, L*4,N,M	A1*3, J, A2*2, L, A3*2, M, H,A5, A6,N*2, F*2	A1*2, B1*3, B2*2,C, E,G,M	C*2, H, L, B2, M

Table 11
Analysis of Writing Test in NMET

Year	2016(I)
Length	112
Coherence word	3
Sentence patterns	L*2,C

Seeing from Table 9, the length of possible version has at least 77 words in 2008 and 152 words at most in 2013. Even though the requirement of basic writing asks for only five sentences and doesn't have the word requirement, we can say that using more compound sentences will be better to help write long sentences so that to build up a solid foundation to express clearly and fully. As for the summary of task-based writing, the requirement asks for 30 words, but actually the words number of summary in 2007, 2008 and 2013 are far more than required. The same thing happened in the body part of task-based writing. It seems that the word limit is not a scoring point. The use of coherence words shows the great importance of the coherence of the discourse. Meanwhile, the wide use of different sentence patterns is good evidence to show the language use ability. Moreover, many long sentences appeared in possible versions, even up to 50 words in one sentence. It is a big challenge for high school students and gives students a hint: pay attention to the sentence patterns with diversity and complexity. As is shown in Table 11, the length of writing task in NMET accords with the requirement and long sentences are seldom used. Therefore, writing task in NMET mainly focuses on conveying information in a logic and coherent way regardless of the sentence complexity.

Table 12
Data Analysis of Candidates' Performance of Basic Writing[5]

Year	Average score of paper	Average score of BW	Percen-tage	Facility index	Difficulty degree	Discrimination index	Standard deviation
2007	74.74	6.74	9.0%	0.38	difficult	0.83	3.73
2008	69.84	6.83	9.8%	0.46	Ok	0.84	3.80
2009	77.36	7.07	9.1%	0.47	Ok	0.84	3.48
2010	91.21	7.04	7.7%	0.47	Ok	0.86	3.52
2011	71.72	6.64	9.3%	0.44	Ok	0.83	3.36
2012	73.31	6.97	9.5%	0.46	Ok	0.86	3.64
2013	74.80	7.04	9.4%	0.47	Ok	0.89	3.85
2014	72.50	7.34	10.1%	0.49	Ok	0.88	3.98
2015	70.57	6.47	9.2%	0.43	Ok	0.88	3.89

Table 13
Data Analysis of Candidates' Performance of Task-Based Writing

Year	Average score of paper	Average score of TW	Percen-tage	Facility index	Difficulty degree	Discrimination index	Standard deviation
2007	74.74	8.95	12.0%	0.36	difficult	0.82	5.78
2008	69.84	11.30	16.2%	0.45	Ok	0.85	5.75
2009	77.36	10.25	13.2%	0.41	Ok	0.85	5.81
2010	91.21	10.74	11.8%	0.43	Ok	0.83	5.99
2011	71.72	12.22	17.0%	0.49	Ok	0.81	5.83
2012	73.31	11.94	16.3%	0.48	Ok	0.88	5.96
2013	74.80	10.86	14.5%	0.43	Ok	0.91	6.3
2014	72.50	11.39	15.7%	0.46	Ok	0.90	6.17
2015	70.57	12.11	17.2%	0.48	Ok	0.90	6.38

[5] The data is quoted from the annual report of GDNMET from 2007 to 2015. The report of 2016 doesn't contain individual data about writing item.

According to the statistics of Annual Report of GDNMET 2007, the facility index is less than 0.4, which means basic writing and task-based writing are too difficult for the candidates, so they didn't get the satisfying result (Alderson et al., 2000). The coming seven tests were getting better in improving the content validity since GDNMET 2007 was the first entrance exam in new curriculum reform.

The facility index shows the difficulty of both two items is moderate except the test in 2007. Candidates' performance maintains basically at 0.41-0.49 level, maybe the reason is that candidates took the new-type items for the first time. Suitable difficulty makes the high-level students have enough space to carry on the display development, and the middle-level students may complete the task-based writing smoothly, while the low-level students can also write some contents, so it won't lose points completely.

In terms of scoring ratio, the writing score rate is low. Basic writing covers 10% and task-based writing 16.7%. According to the above tables, basic writing does a good job except 2010; task-based writing need to be improved, since the covering varies year by year from the scale 11.8% to 17.2%. As is shown in the tables, the discrimination index and standard deviation tell us that the score of test items are valid.

From analyzing the reading materials, possible versions and data of candidates' performance, we can find that the writing content is suitable for students.

CONCLUSION

According to the analysis above, it is clear to find that in new curriculum reform, the contents of basic writing and task-based writing tests from 2007 to 2014 conform to the requirements of ECSCSMS and testing syllabus, which is suitable to most of the candidates as well. What's more, two testing items are representative to measure students' writing ability. It is evidently proved that the writing tests are sufficiently representative and validity of what writing is supposed to test the candidates for entrance examination, so they have the high content validity.

IMPLICATION

There are some implications for the trend of writing test in the coming testing reform: a) Given a situation, table, or image will be more often used to let the examinees develop a bigger space in writing; b) topics will still be closely related to the social hot spots and associated with the real life of middle school students; c) the weighting of writing testing will not reduce less than 40 points; d) the two testing items will not reduce to one.

Meanwhile, some teaching enlightenments for high school English teaching are available. Teachers should pay attention to the basic writing and cultivate the writing ability of students and let the students learn to use basic sentences to express themselves rather than to remember some formulas to deal with the task. Ask the examinees to understand the complex sentence, to learn how to use the complex sentence to express information points, try to avoid some mistakes for they are easily affected by the way of Chinese thinking and writing. Writing topics are widely selected, most of which are related to people's social life, while some of them are current hot topics. As a result, teaching should be expanded to extra-curricular activities. English teachers should strengthen the extensive reading and broaden the resources of learning English, such as reading English newspapers or English abbreviations novels, listening to English news and learning English online, etc.

REFERENCES

Alderson, J. C. (2000). *Assessing reading*. Cambridge: Cambridge University Press.

Alderson, J. C., Clapham, C., & Wall, D. (2000). *Language test construction and evaluation.* Beijing: Foreign Language Teaching and Research Press.

Bachman, L. F. (1990). *Fundamental considerations in language testing.* Oxford: Oxford University Press.

Heaton, J. B. (2000). *Writing English language tests (new edition).* Beijing: Foreign Language Teaching and Research Press.

Henning, G. (2001). *A guide to language testing: Development, evaluation and research*. Beijing: Foreign Language Teaching and Research Press.

Liu, R. Q., & Han, B. C. (2000). *Language testing and technique.* Beijing: Foreign Language Teaching and Research Press.

Wood, R. (2001). *Assessment and testing: A survey of research*. Beijing: Foreign Language Teaching and Research Press.