A Preliminary Study on the Validity of English Passage Dictation Test[*]

PENG Mei

Shanghai International Studies University, Shanghai, China

This paper tries to demonstrate that passage dictation test is a valid testing method to test learners' comprehensive English ability via a correlation study with the supplement of a questionnaire survey and interviews. Due to the fact that TEM (Test for English Majors) tests have been proved to be reasonably reliable and valid in China and are commonly acknowledged as well-established English proficiency language tests, we therefore attempted to conduct a study in which we compared the learners' passage dictation test scores with their total scores and scores of each parts of TEM4 (TEM-Band Four) with the help of a statistical software (SPSS (Statistical Program for Social Sciences)) to see if they were correlated. The results of the correlation study (the correlation coefficient between passage dictation test scores and the total scores is 0.756 at the 0.01 level, which means highly significant) together with the questionnaire and interview results provided fairly solid evidence that passage dictation test is indeed a valid and reliable testing method to assess test takers' comprehensive English competence.

Keywords: English passage dictation, validity, correlation study

Introduction

Dictation has been widely employed in foreign/second language teaching in that it is not only a useful and effective teaching technique but a practical and easy testing method (YANG, 2003, 2009; ZHANG, 2004). Quite a few scholars (Oller, 1979; Coniam, 1995; Buck, 2001; Alkire, 2002; Smith, 2010) have pointed out that the process of dictation reflects nearly all aspects in language communication, which involves phonological, grammatical, and lexical knowledge, and that listeners need to use their analytic competence in the discourse level while having dictation practice which meanwhile can train learners' listening ability, spelling, and short-term memory capacity, and conclude that dictation is a positive and creative psychological skill training. In China, dictation tests are frequently seen in nationwide tests such as Test for English Majors (TEM4 and TEM8) and College English Test (CET4 and CET6). Some scholars (Oller, 1979; KONG & NIE, 2002) classified dictation tests into the following types: (1) passage dictation test (also called traditional dictation in which learners are asked to write down everything in a passage they hear four times); (2) spot dictation test (also called partial dictation in which learners write down the missing words in the blanks while listening to a passage); (3) compound dictation test (in which learners are asked to write down missing words and a few sentences in the blanks while listening to a passage); and (4) dicto-com (also called reproduction in which learners work in groups to recreate a passage they have just hear by taking notes). This paper focuses on the

discussion of the validity verification of English passage dictation test by conducting a correlation study.

Theories About Validity

Validity is described in the Standards for Education and Psychological Testing as "the most fundamental consideration in developing and evaluating tests" which refers to "the degree to which evidence and theory support the interpretations of the test scores entailed by proposed users of tests" (JIN & FAN, 2011). In fact, since the end of the 1970s, validity began to be viewed as a unitary concept which was construct-based. Bachman and Palmer (1996) thought that the term construct validity is used to refer to "the extent to which we can interpret a given test score as an indicator of the ability(ies), or construct(s) we want to measure" (p. 21). In other words, it pertains to "the meaningfulness and appropriateness of the interpretations that we make on the basis of test scores" (Bachman & Palmer, 1996, p. 21). Many scholars reached the consensus that construct validity "is the whole of validity" (Loevinger, 1957, p. 636) and it was seen as a general approach to validity that includes all evidence for validity, including content and criterion evidence, reliability, and the wide range of methods associated with theory testing (Messick, 1975).

According to Bachman and Palmer (1996), when we interpret scores from language tests as indicators of test takers' language ability, a very important question involves how we can justify the interpretations or demonstrates the validity of the interpretations instead of simply asserting or arguing that they are valid. In other words, we need to provide evidence that the test score reflects the area(s) of language ability we intend to measure. To provide such evidence, we must first define the construct we want to measure, that is, the specific definition of an ability that provides the basis for a given test or test task and for interpreting scores derived from this task. When it comes to passage dictation test, we believe the construct underlying it is the test takers' comprehensive language competence, which mainly consists of learners' listening comprehension level, knowledge of grammar and vocabulary, writing ability and knowledge of the target language culture (Oller, 1979; Brown & Barnard, 1975; Harrison, 1983; Buck, 2001; YANG, 2003).

Research Methodology

Research Design and Research Questions

According to JIN and FAN (2011), TEM tests have been proved to be "reasonably reliable and valid tests that are set at an appropriate (difficulty) level as defined in the test specification" (p. 593), and as a matter of fact, it is commonly acknowledged that TEM has grown into a well-established English proficiency language test in China. In TEM4, passage dictation accounts for 15% of the total test score. Therefore, we attempted to justify the construct validity of passage dictation test via a correlation study. The underlying idea of the study is: If the correlation coefficient is high between passage dictation test scores and the total scores of TEM4, we can basically claim that passage dictation test is valid in assessing test takers' comprehensive English ability. Based on this hypothesis, we posed the following research questions: (1) How are test takers' passage dictation test scores correlated with their total test scores of TEM4?; and (2) Is passage dictation test a valid testing method in testing test takers' comprehensive English competence?.

Participants and Setting

The participants in this study were 38 first-year adult English learners majoring in British and American studies in the Continuing Education College of Shanghai International Studies University. This is a three-year program in which the students attend to the night school three evenings a week and one Saturday afternoon in

the hope of obtaining a bachelor's degree at the end of the program. The participants had passed the College Entrance Exams for Adult Learners held by the Education Assessment Institution of Shanghai to get admitted. Before entering this college, they had learned English for about 8-11 years. The author taught the Comprehensive English Course to them twice a week. Before the study was carried out, the students had been receiving in-class English passage dictation practice once a week for about 20 minutes for almost a whole academic year. Altogether, they had 25 passage dictation practices with the help of a tape recorder, and all the passages and audio recordings were chosen from published TEM4 tests guides to make sure that the students got standard and authentic passage dictation practice to the maximum extent. Every time they had dictation practice, keys and standard rating requirements for grading the passage dictation would be released to them through online chat room shortly after the class was over.

Research Methods and Instruments

This study was descriptive in nature, and we mainly made use of the method of correlational research to investigate, simultaneously, the relationship between two variables, namely, passage dictation test score and total test score of TEM4, and the relationship between passage dictation score and other test items scores in TEM4 including listening comprehension, grammar and vocabulary, cloze，reading comprehension, and composition.

The instruments in this study consist of the following things: (1) a simulation test of TEM4—"Sample Test 1" from ZOU (2005), in which the reliability and validity can be guaranteed; (2) a questionnaire designed by the author; (3) interviews; and (4) the teacher's observation.

Data Collection Procedures

Firstly, the participants were arranged to take part in the simulation test of TEM4 at the end of their first academic year. Secondly, all the test papers were graded. The MCQ (Multiple Choices Questions) items were rated by the author according to the keys provided by the sample test, while a colleague was invited to double rate the passage dictation part and compositions in accordance with the standard rating requirements for passage dictation and writing. The rater consistency is above 0.90 ($r > 0.90$). Thirdly, all data (scores of each test item and the total scores of the simulation test of TEM4) collected were entered onto SPSS 13.0 (Statistical Program for Social Sciences) and analyzed for mean frequencies and Pearson coefficient.

The following steps were taken in order to collect qualitative evidence: (1) Shortly after the simulation test, all the participants were arranged to fill in a questionnaire in which they were asked to reflect on the process of passage dictation test, including what sub-abilities or language aspects they thought were needed in passage dictation, and what the biggest difficulty/challenge was for them and if they thought their performance in passage dictation was basically an indicator of their overall English competence. All the 38 participants responded the questionnaire which turned out to be all valid; and (2) The author interviewed 12 students with different English levels based on their simulation test scores and their performance at school (four good learners, four average learners, and four poor learners) to ask them to express their views on the author's questions related to passage dictation.

Data Analysis and Discussion

Statistical Analyses of the Simulation Test Results

Table 1 shows the descriptive statistics of each test item in the simulation test. As we can see, the mean of passage dictation is 8.66 (the full mark is 15 points) with the highest score 14.50 and the lowest 2.50. The mean

of the total score is just 54.04 (the full mark is 100 points) with the highest score 74.5 and the lowest 37.00. This indicates that the participants are generally lower-intermediate English learners and their levels of English are quite uneven.

Table 1

Descriptive Statistics of the Test Items in the Simulation Test

Items	N	Mean	Std. deviation	Max.	Min.
Dictation	38	8.66	3.150	14.50	2.50
Listening comprehension	38	8.74	1.560	12.50	5.50
Grammar and vocabulary	38	7.28	2.770	14.00	3.00
Cloze	38	4.00	2.175	10.00	1.00
Reading comprehension	38	9.26	2.583	15.00	5.00
Composition	38	16.33	2.060	20.50	10.00
Total score	38	54.04	9.540	74.50	37.00

The author used Pearson coefficient to check out the correlation between scores of passage dictation and those other test items as well as the test totals. Table 2 displays the correlative results. We can see that the correlative coefficients for the relationship between passage dictation and listening comprehension, cloze, and test total were highly statistically significant ($p < 0.01$), while passage dictation and writing are correlated at the 0.05 level ($p < 0.05$). The results basically support our assumption that people who do well in passage dictation test also perform well in the whole test (Pearson $r = 0.756$), and learners' performance in passage dictation is highly correlated with their performance in listening comprehension (Pearson $r = 0.632$) and writing (Pearson $r = 0.326$), which also validates the research findings of similar studies (LI, 2001; ZHANG, 2004). As it is known to us, cloze test is also regarded as a testing method assessing test takers' comprehensive language ability, there is no wonder that passage dictation and cloze were significantly correlated. However, we find passage dictation is not significantly correlated with grammar and vocabulary, and reading comprehension. Interestingly, this was against our previous assumption that passage dictation should be highly correlated with grammar and vocabulary. With the help of questionnaire and interviews, the author found the possible reasons: The participants were averagely weak in grammar and were quite poor at making use of their learned grammar knowledge to correct or avoid their mistakes during the dictation process. As to vocabulary, they think passage dictation put a relatively low demand for their range of vocabulary and their ability to use a particular word exactly in a suitable circumstance. What they needed was actually to just write down correctly the words they heard, which rather required their ability to spell words accurately than to use a word creatively or appropriately.

Table 2

A Correlational Study of Passage Dictation and the Simulation Test of TEM4

Items	N	Pearson correlation	Sig. (2-tailed)
Dictation & listening comprehension	38	0.632**	0.000
Dictation & grammar and vocabulary	38	0.218	0.188
Dictation & composition	38	0.326*	0.046
Dictation & cloze	38	0.437**	0.006
Dictation & reading comprehension	38	0.294	0.073
Dictation & total score	38	0.756**	0.000

Notes. ** Correlation is significant at the 0.01 level (2-tailed); * Correlation is significant at the 0.05 level (2-tailed).

Analysis of the Questionnaire

Table 3 shows the items and results of the questionnaire the participants responded shortly after the simulation test. According to their answers, the participants thought passage dictation test could assess their abilities or knowledge of listening, phonetics, spelling, grammar and vocabulary, and writing. They thought the greatest difficulty/challenge for them in passage dictation came from (according to the order of difficulty) low listening level (38%), limited range of vocabulary (35%), poor spelling ability (19%) and weak grammar foundation (8%). Seventy-nine percent of the participants agreed that their performance in passage dictation was basically predictor of their comprehensive English competence. In general, the author found the results of the questionnaire were useful additional evidence to support the construct validity of passage dictation test.

Table 3

Questionnaire Results (n = 38)

Q 1: What aspects of language do you think is/are assessed in passage dictation test? (more than one choice is allowed)					
Listening	Phonetics	Grammar & vocabulary	Spelling	Writing	All above mentioned
68%	49%	14%	41%	5%	19%

Q 2: What do you think is the biggest difficulty (challenge) in the process of taking passage dictation test? (one choice only)			
Poor listening ability	Vocabulary	Grammar	Spelling
38%	35%	8%	19%

Q 3: Do you think your performance in passage dictation is basically an indicator of your comprehensive English competence?				
Strongly agree	Agree	Not sure	Disagree	Strongly disagree
28%	51%	15%	6%	0

Analysis of the Interview

Among the 12 students who were interviewed, all admitted that passage dictation could assess their comprehensive English ability, especially their listening ability, spelling, vocabulary, and grammar knowledge. Three average learners believed dictation tested more listening than other aspects of language. For instance, one student said to the author:

> If I can't understand what I hear, I can write down nothing, even if I have actually learned the words before and have no difficulty spelling them. I think the reason why I did unsatisfactorily in passage dictation is mainly because my listening is poor and I can't recognize those words or sentences with my ear. (Personal information, June 15, 2010)

However, they all agree that regular and constant passage dictation practice is a very good and effective way to improve their listening and comprehensive English ability, and that they all benefited a lot for the in-class passage dictation practice in the past academic year.

Teacher's Observation and Evaluation

Having taught the participants for a whole academic year, the author found effective learners almost always did well in passage dictation quizzes or test, and those who performed well in passage dictation tests were invariably good language learners. The author believes her observation could also contribute to the evidence that learners' performance on passage dictation can normally predict or indicate their comprehensive English competence in foreign/second language learning.

Conclusions and Suggestions

This paper tries to demonstrate that passage dictation test is a valid testing method to test learners'

comprehensive English ability via a correlation study with the supplement of a questionnaire survey and interviews. The assumption is made that English learners who do well in passage dictation test will also perform well in the whole proficiency test, and learners' performance in passage dictation may be highly correlated with their performance in listening comprehension, grammar and vocabulary, and writing. To a large extent, the findings of the correlation study together with the questionnaire and interview results supported our assumption. We can basically come to the conclusion that passage dictation test is a quite valid testing method to assess test takers' comprehensive English competence and it is a practical indicator of learners' language proficiency level.

However, the author would also like to make some suggestions to the Chinese NACFLT (National Advisory Committee for Foreign Language Teaching) based on her teaching experience concerning passage dictation in hope of perfecting this testing method. First, the author thinks we had better pay more attention to the accent and talking speed of the native speakers who are invited to record the reading of passage dictation for TEM4. Quite a few teachers including herself have discovered that sometimes the speaker's accent is so strong that it is quite hard to recognize the words he/she speaks. And the speakers' talking speed is not very consistent, either. The author found they speak faster in some years than in other years. Second, the author thinks the objectivity of scores of passage dictation in TEM4 might be better achieved if it could be double rated. The validity of passage dictation test may be enhanced if the above-mentioned problems are given more attention and can be satisfactorily solved.

References

Alkire, S. (2002). Dictation as a language learning device. *The Internet TESL Journal, VIII*(3). Retrieved from http://iteslj.org/Techniques/Alkire-Dictation. html

Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice*. Oxford: Oxford University Press.

Brown, D., & Barnard, H. (1975). Dictation as a learning experience. *RELC Journal, 6*(2).

Buck, G. (2001). *Assessing listening*. Cambridge: Cambridge University Press.

Coniam, D. (1995). Computerized dictation for assessing listening proficiency. *CALICO Journal, 13*(2-3).

Harrison, A. (1983). *A language testing handbook*. London: Macmillan Press.

JIN, Y., & FAN, J. S. (2011). Test for English majors (TEM) in China. *Language Testing, 28*(4), 589-596.

KONG, Y. P., & NIE, J, Z. (2002). Compound dictation in CET and its backwash effect on language teaching. *Foreign Language World,* (2).

LI, S. S. (2001). Dictation: An effective means of FL teaching and teaching evaluation. *Journal of PLA University of Foreign Languages,* (4).

Loevinger, J. (1957). Objective tests as instruments of psychological theory. *Psychological Reports, Monograph Supplement, 3,* 635-694.

Messick, S. (1975). The standard program meaning and values in measurement and evaluation. *American Psychologist, 30,* 955-966.

Oller, J. (1979). *Language tests at school*. London: Longman.

Smith, D. (2010). Pair-dictation activities for teaching English to university students. *The Internet TESL Journal, XVI*(2). Retrieved from http://iteslj.org/Techniques / Smith-Dictation. html

YANG, M. Z. (2003). The effect of dictation practice on the language skills of Chinese EFL learners. *Modern Foreign Languages,* (3).

YANG, M. Z. (2009). A study of the effects of dictation methods on dictation test results. *Foreign Language Learning Theory and Practice,* (2).

ZHANG, J. Q. (2004). Research on strengthening dictation practices in English teaching. *Media in Foreign Language Instruction,* (2).

ZOU, S. (Ed.). (2005). *A guided workbook for English majors (Grade 4)* (2nd ed.). Shanghai: Shanghai Foreign Language Education Press.