

An evaluation on the validity of English tests used for English 10 at some Higher Secondary School in the Middle and North of Vietnam, from Ha Tinh to Ha Nam = Đánh giá tính giá trị của các bài kiểm tra tiếng Anh dành cho học sinh lớp 10 ở một số trường THPT miền Trung và miền Bắc Việt Nam, từ Hà Tĩnh đến Hà Nam. M.A Thesis Linguistics: 60 14 10 \ Nguyễn Thị Hoàng Liên ;

Supervisor : Dr. Hà Cẩm Tâm

The primary objective of this thesis is to evaluate the validity of English tests used for English 10 at some Higher Secondary Schools in the Middle and North of Viet Nam, from Ha Tinh to Ha Nam. Specifically, it aims at finding out whether the collected tests are valid in term of content, construct and face validity. The author is only interested in language components in the collected tests. The specific objectives of the thesis are outlined as follows:

- To investigate whether the testing techniques used in the tests correctly measure pupil's ability to remember, understand and produce the learnt language components.
- To investigated whether the language components in the tests are a good representation of the syllabus.

- To determine whether the language components tested truly reflect what pupils were taught.

- To give some implications to improve the collected tests which are invalid

A combination of both qualitative and quantitative methodologies was used to achieve the objectives of the thesis.

The author was encouraged to conduct this study for the following reasons:

First, in the recent years there have been a lot of changes involving teaching methods, teaching materials as well as testing at Higher Secondary School in Viet Nam. However, there have been no studies to evaluate whether the tests used at Higher Secondary Schools test what pupils are taught.

Seconds, the thesis was conducted on validity because this test characteristic is critical for tests and referred to as integral measurement qualities. Because

this quality provides major justification for using test score numbers as a basic for making inferences or decisions (Bachman and Palmer, 1996:19). Moreover, the emphasis on test validity is also confirmed in Hughes (1989) that, "the greater a test's content validity is, the more likely it is to be an accurate measure of what it is to measure." Bachman (1990: 289) also insists that : "The most important quality to consider in the development, interpretation and use of language tests is validity, which has been described as a unitary concept related to the adequacy and appropriateness of the way we interpret and use test scores."

The last reason lies in the testing situation at High Secondary Schools in the Middle and North of Viet Nam. Neither comprehensive nor systematic evaluation nor research on the effectiveness and appropriateness of the achievement tests at High Secondary Schools under investigation. No formal discussions or seminar on test construction or test methods have been carried out. In addition, There is a

lack of a language test item bank, a professionals testing committee, who judges the quality of the tests and takes the responsibility for the given tests.

The thesis is organized into three chapters.

In the chapter one, the thesis embarks on basic concepts of testing, reviews achievement tests because the collected tests are of progress and final achievement tests, provides characteristics of a good test focusing on validity with an emphasis on construct, content, face validity and source of invalidity. Test items for phonetics, structures and vocabulary and curricula including the objectives of the 10th form course and the teaching material are also presented in chapter one.

Chapter 2 looks to the study which is the main part of the thesis. One main question following three minor questions are presented as follows:

Q1: *Do the achievement tests for Higher Secondary School pupils meet the following criteria: content validity, construct validity, face validity?*

Question 1 is further specified by minor questions evaluating on each type of validity (construct, content and face validity) respectively as follows:

Q 1.1: *Can the testing techniques used in the tests correctly measure pupils' ability to remember, understand and produce the learnt language components?*

Q 1.2. *Are the language components in the tests a good representation of the syllabus?*

Q 1.3. *Do the language components tested truly reflect what pupils were taught?*

Then the data are described. The data of the thesis are progress and final achievement tests collected from some High Secondary Schools in the Middle and North of Viet Nam from Ha Tinh to Ha Nam. Thirty achievement tests taken from the schools under investigation include twenty progress achievement tests (ten fourty five minute written tests and ten fifteen minute written tests) and ten final progress achievement tests (the first or second

semester tests). The forty five written tests which are administered at the end of three or two lessons serve as testing pupils' reading, listening, writing skills, phonetics, structures and vocabulary. The term and final term tests are usually structured like forty five minute tests. This kind of tests is administered to test grammar, vocabulary and reading, writing, listening skills relating to the themes that pupils learn during the term or the school year.

The fifteen minute tests , however only consist of structures or vocabulary or both and phonetics that pupils learn during the week or in the previous lesson.

In order to find out whether the available English tests designed and used at Higher Secondary Schools under investigation are in accordance with the syllabus and objectives of the 10th form or not, both objective and subjective methods which mainly based on the analysis of random samples of collected progress achievement tests and final achievement tests at the given Higher Secondary Schools.

The author based on the definition of "construct validity of a test" by Hughes, A, 1989:26 to evaluate the construct validity of thirty tests collected. The construct validity will be measured through testing techniques used. If the testing techniques used can correctly measure the language components which are intended to be tested in the tests, the employed testing techniques are said to possess construct validity. The author also analyzed whether the testing technique can check pupils' ability to remember, understand, use and produce the learnt language components. The weaknesses and strengths of each technique will be evaluated to draw the conclusion. Because of time constraint, the author didn't carry out the experimental study. Basing on the comparison between the test specification and the proposed test specification of language components, it was found out that six of thirty collected tests (20%) lack construct validity because teachers only employed multiple choice questions techniques to measure pupils' Phonetics, structures and vocabulary.

Although this technique is effective to a wide range of knowledge, it sometimes fails to reflect real ability of pupils because pupils tend to choose the correct answers by chance when they are unable to decide which one is correct. In addition, this test technique may be a poor indicator of someone's ability to use structures and vocabulary because the person who can identify the correct response in the item above may not be able to produce the correct one when writing or speaking. The rest tests possess construct validity in the meaning that the test items used can test pupils' ability to distinguish learnt sounds, use learnt structures appropriately and use vocabulary in different contexts.

To investigate the content validity, the author analyzed and compared the tests' content (language components: phonetics, vocabulary and grammar) with the proposed test specification basing on the objectives and syllabus of the textbook. From the comparison between the content of the tests collected and the content required to be tested in the textbook,

the author found out that ninety (90%) of the tests analyzed contain knowledge of phonetics, structures and vocabulary listed in the specifications. To put in another way, these achievement tests were designed basing on the textbook- English 10 and the detailed course syllabus. In addition, they are also based on what the pupils have already learnt in the course book and the level of the pupils. Therefore, it can be concluded that these collected tests are valid in terms of content. The other test (10%) lack content validity because they test what are not included in the objectives of the syllabus.

The author designed the survey questionnaire comprising of three questions with an aim at investigating teachers' reflection of the test validity. The questionnaire was written in English for teachers and was given to 30 teachers at schools under investigation. All of the returned forms were valid. After analyzing the result of the survey questionnaire, the author realized that most tests ensure face validity in the meaning that phonetics, structures and

vocabulary tested can reflect what pupils are taught. However, these parts cannot truly measure what pupils have learnt because of the testing techniques used (multiple choice questions). Pupils can choose correct answer by chance.

By employing both quantitative and qualitative method, the author draw a conclusion that In conclusion, most of the collected tests possess construct and content validity; only a few tests lack construct validity, and content validity. Besides, the contents tested were not distributed evenly between vocabulary and structures. Some tests contain too many structure items, some tests include too many vocabulary items. All of the teachers interviewed agreed that the tests collected in general can reflect what pupils were taught.

The last chapter provides the summary; suggestion for improving the invalid tests. First, to improve construct validity of the tests under investigation, two or three testing techniques should

be employed to measure pupils' ability of using and producing structures and vocabulary. For example, to test vocabulary, both multiple choice question and word formation technique should be employed. Besides, other techniques are used flexibly such as word formation, word form, sentence transformation and sentence building. Both objective and subjective method should be used to ensure that teachers can give the most accurate as possible picture of pupils' ability. Second, teacher can improve all the tests that are invalid in term of content validity because of testing stress which is not included in the objectives of the syllabus by omitting stress items.

The two given solutions can solve the problems in the ten invalid tests. However, more attention should be paid to test designing to ensure construct, content and face validity as follows:

First, to improve construct validity of the tests under investigation, two or three testing techniques

should be employed to measure pupils' ability of using and producing structures and vocabulary.

Second, teacher can improve all the tests that are invalid in term of content validity because of testing stress which is not included in the objectives of the syllabus by omitting stress items.

Third, both construct, content and face validity of the invalid collected tests can be improved if teachers inspect test items regarding their correspondence to the objectivities of the textbook. The test writer should clarify clearly the test objectives prior to test development. Then drawing up a detailed specification for the test is of great importance to guarantee the test objectives and therefore to ensure that the test possess content validity. This can be done by selecting and constructing the items carefully basing on the syllabus of the textbook. It is also advisable that teachers well prepare the content of the test before designing tests. They shouldn't take ready-made tests elsewhere because that do not truly reflect the teaching content

and thus disadvantage pupils. Therefore, teachers ought to well select authentic materials for test inputs in order that the test can reflect the goals of teaching and is closely related to what pupils are taught and to the textbook content.

Like any other research, this research is working under limited conditions and has some limitations.

First, sample achievement tests have not yet developed, piloted and evaluated due to constraints of finance and administration. Second the test evaluation is limited to evaluating content, construct and face validity, other factors have not yet been taken into consideration such as practicality, concurrent validity, reliability, etc.

Besides, the study only focuses on the language components of the tests, other areas such as skills: reading, writing or listening were not mentioned.

Third, the backwash effect of the achievement tests has not been paid attention to due to limited time.

Fourth, to evaluate the face validity of the tests, the author only administered questionnaire to teachers,. The student population was not included because of time limitation and some other reasons relating to the time lap between the tests and the study.

The study has several suggestions for further future research work to generate a more comprehensive picture of achievement tests used for pupils at Upper Secondary Schools from Ha Tinh to Ha Nam. First, the backwash of the test should be evaluated. Second, the reliability of the tests should be also determined. Last, the problems with test development ought to be found out.

To sum up, in spite of its small scale, this thesis may make a positive contribution to evaluating not

only the current English achievement tests used for the 10 form pupils in terms of content, construct and face validity but also the language testing situation at Upper Secondary Schools from Ha Tinh to Ha Nam. It is hoped that the dissertation will improve worthwhile to those who are involved in the design of their own tests and encourage other researchers to engage more in the field, especially in those areas identified for future research. It is also wished that this dissertation will greatly contribute to the improvement of language testing at Upper Secondary Schools in the North and Middle of Viet Nam from Ha Tinh to Ha Nam.