

An Exploration into Teachers' Perceptions in Designing English Achievement Tests at Primary Schools in Thai Nguyen City, Vietnam



Dinh Thi Lien

School of Foreign Languages, Thai Nguyen University

ABSTRACT: Language testing and assessment has a direct impact on teaching because teachers utilize test results to modify their course objectives, make better use of the materials and activities in class, and, if they are the testers, design better tests. However, there often have been a lot of gathered concerns in relation to the content validity of locally designed achievement English tests from the educators and administrators who always raise a question of whether the teacher-made English achievement tests faithfully reflect the course objectives, expected linguistic knowledge, skills as well as measures the real linguistic competence of test-takers. This study aims to seek out the evidence to comprehend the reality of teacher-made English achievement tests and provide them with the required teacher development programs to help them improve their proficiency in teaching and testing. The results show the majority of teachers have strong perceptions about EAT and have a good understanding of testing principles; therefore, they conducted their tests using an appropriate process while keeping in mind the English teaching and testing context. It can be noted that the findings from the study have shed light on important factors which relate to the effect of the process of test preparation and test construction.

KEY WORDS: language testing, teachers' perception, test design, achievement test

I. INTRODUCTION

It is widely accepted that English has now become the international lingua franca and is playing a more and more dominant role in people's lives, ranging from politics, economics, and health care, to entertainment and education. More specifically, many countries have changed their educational syllabi so that children are starting to learn English at a younger age.

In line with globalization, English language teaching and learning is a critical issue in educational systems around the world, including in developing countries like Vietnam where English is considered as a foreign language. In Vietnam, learners' acquisition of English can be dramatically influenced by government policy toward the role of English in society. Under these circumstances, English is taught as a compulsory subject in primary, junior, and senior high schools and as a complementary subject in universities and colleges. As a direct result of this development, it is now necessary to teach and assess students' English proficiency at the institutional level.

According to Brown (1996), language tests are used to gather information about the learner's language abilities in order to make informed educational decisions. Thus, it is important to have good English tests to be able to make good decisions on learners' English proficiency. In the same view, Madsen (1983) holds the view that "good language tests help students learn the language by requiring them to study hard, emphasizing course objectives and showing them where they need to improve" (p.5). However, what makes a good test so that language teachers can make decisions about their classrooms based on the data it provides? Several researchers indicate that the characteristics of a good test must be valid, reliable, objective, and practical (Bachman, 1990; Bachman & Palmer, 1996; Brown, 1996; Henning, 1998).

Achievement tests are those that provide direct feedback to teachers and students. Teachers and students can use achievement tests to determine "the amount of language that each person is learning in a given period of time" (Brown, 1995, p.111). Thus, they support both parties in making judgments on the teaching-learning process. Brown (1996) also points out that the achievement tests used in a language program may be adopted, developed, or invented. Heaton suggests that "the best tests for the classroom are those you write yourself" (1990, p.23) because the teacher knows what learners need, what subjects have been covered in class, and how, which will help to maintain high content validity. Furthermore, the language in the classroom can be seen as a distinct social context that is specially designed for educational purposes to encourage language learning. Obviously, teachers play a crucial role because they are the last users who implement the curriculum, have a thorough understanding of their students, and constantly engage with the students to support and motivate them to advance as planned in the school curriculum. As a result, all

An Exploration into Teachers' Perceptions in Designing English Achievement Tests at Primary Schools in Thai Nguyen City, Vietnam

language teachers in the classrooms frequently conduct tests, not only to evaluate students' development and performance but also to guide them in furthering their studies.

On the contrary, there often have been a lot of gathered concerns in relation to the content validity of locally designed achievement English tests from the educators and administrators who always raise a question of whether the teacher-made English achievement tests faithfully reflect the course objectives, expected linguistic knowledge, skills as well as measures the real linguistic competence of test-takers. Similar concerns also occur to the case of the teacher-made English achievement tests at a primary school in Thai Nguyen city, Vietnam. In the light of these concerns, the researcher has been inspired and motivated to conduct the research on the teachers' perceptions and practices with the title "*An exploration into teachers' perceptions in designing English achievement tests at primary schools in Thai Nguyen city*" with the belief that they will comprehend the reality of teacher-made English achievement tests and provide them with the required teacher development programs to help them improve their proficiency in teaching and testing. It might be assumed that, as a result of this procedure, teachers will have the opportunity to reflect on why they do and what they do when testing their students.

II. LITERATURE REVIEW

2.1. Definition of language testing

In fact, there are many methods for assessing the teaching-learning process. One method is to conduct a test. Test is a way to measure the knowledgeability of a person or student. According to Brown, test is a method of measuring a person's abilities, knowledge or performance. In line with Brown's opinion, Sadar (2020) said that a test may be a tool, a question, a set of questions, or an examination that is used to measure a specific feature of an individual or a group of individuals. Penny (1996) said that "Tests are used as a means to motivate students to learn or review specific material" (p. 34). It means test is one motivation for students to learn or review material in their school. Furthermore, Fernandes (1984) states that a test as a systematic procedure for surveying a person's behavior and explaining it with the aid of a numerical scale or a category system.

In addition, according to Linn and Gronlund (1995), "a test is a particular type of assessment that typically consists of a set of questions administered during a fixed period of time under reasonably comparable conditions for all students" (p.5). Globally, the test consists of a series of questions that must be answered correctly, with the purpose of determining an individual's ability to understand everyday learning.

Testing is fundamental in any language program. According to Brown's (1995) model of evaluating the elements of the language curriculum, since from the placement of students in the program through their graduation, all kinds of program-related decisions are made using the tests as a basis.

Based on the aforementioned definitions, the researcher concludes that a test is a specific sort of evaluation used to reinforce learning and motivate students by assigning a task or a group of activities. The teacher uses the test to not only evaluate and encourage pupils but also to enhance the lesson in the teaching and learning process. To make an accurate decision, the teacher requires reliable data and the ability to gather data, which necessitates the use of a good instrument.

2.2. Relationship of teaching, learning and testing

Testing is an element of the curriculum. The relationship between tests, teachers, and students is so complex that how testing is done gains great importance. In some institutions each teacher is expected to write their own tests. In some others, teachers working at the same level prepare the tests for their groups together. Yet in some others, teachers do not prepare tests at all.

Teaching, learning and testing become involved and related in a learning process like a cycle. First, a teacher must teach the students a specific subject or matter, whether it is a specific grammatical structure, vocabulary, or even a story. Then, the teacher must ensure that the learners completely understand the material that the teacher is providing. At this stage, the four skills of reading, writing, speaking, and listening must be practiced by the students, and it is the teacher's responsibility to select the tasks that will be most useful to the students' learning. Finally, teachers should make tests. Students should take the assessment provided by the teacher, and the suggestions that were made throughout the course, in order to apply them when the testing comes (Shohamy, 1993).

After this final stage, the cycle begins again, and as teachers, they must teach again, then assess, and finally evaluate. However, at each stage, the teacher should consider some of the most important points, such as the ones mentioned above, in order to improve, simplify, and tailor the learning process to the needs of the learners.

In brief, testing has a direct impact on teaching because teachers utilize test results to modify their course objectives, make better use of the materials and activities in class, and, if they are the testers, design better tests (David, 1990). However, despite its importance, most teachers have little idea of what makes a good test or what qualities a test must possess (Basanta, 1995).

2.3. Achievement tests

Different types of tests are used for different purposes. They assist in the gathering of many types of data. This research concentrates on achievement tests because it aims to discover teachers' perceptions and practices in designing English achievement tests at primary schools in Thai Nguyen city.

An Exploration into Teachers' Perceptions in Designing English Achievement Tests at Primary Schools in Thai Nguyen City, Vietnam

Achievement tests are the tests that teachers are most often involved in constructing, administering, and scoring. These tests are based on the course objectives and indicate how much of them have been met (Hughes, 1989).

For a closer and clearer look, Hughes (1989) and Brown (1996) emphasized on the importance of achievement tests, hence the researcher would like to go over the various types of tests in further detail. A formal examination given at the end of the school year or at the completion of a course is known as an achievement test. The achievement test may be written and administered by ministries of education, officially examining boards, or by members of teaching institutions. In agreement with that, Basanta (1995) asserted that achievement tests play a crucial role in language learning programs. "If we assume that a well-planned course should measure the extent to which students have fulfilled course objectives, then achievement tests are a central part of the learning process."

According to Hughes (1989), achievement tests are closely tied to language courses, with the goal of determining how effective individual students, groups of students, or entire courses have been in meeting objectives. There are two types of achievement tests: final and progress.

"Final achievement tests are administered at the end of a course of study and their content must represent the content of the courses which they are based upon. Progress achievement tests are administered to measure the progress learners are making" (Hughes (1989, p.10-11).

Progress achievement tests can also be used as a diagnostic test to pinpoint students' weak and strong points in learning.

Summative tests or final achievement tests, are attempts to use information about students or programs after a series of educational segments has taken place. Their goal is to provide a summary of how well a certain student, group of pupils, or teacher performed on a set of learning criteria or objectives. Teachers use summative test data to determine grades and to explain reports given to students and their parents. It is assumed that test scores in summative tests have generalizable meaning; that is, the score can be construed to mean something other than the context in which the learner is tested (Glenn, 2010, p. 20).

According to Brown (1996), an achievement test is linked to specific classroom topics, sections, or even the entire curriculum. Achievement tests are limited to certain subjects covered in a curriculum over a particular time period, and they are given after a course has covered the objectives in question. Furthermore, achievement tests are parts of the syllabus, measuring curricular goals (Bachman, 1990; Bailey, 1998; Brown, 2013; Brown & Hudson, 2002; Cheng, Watanabe & Curtis, 2004; Hughes, 2003). Within the framework of this research, achievement tests are evaluated in terms of their role, definition, and types.

In order to create a good achievement test, a test designer should keep in mind that achievement tests must be well-constructed by adhering to the following essential principles (Norman, 1968, p. 4)

- Achievement tests should measure clearly defined learning outcomes that are in harmony with the instructional objectives.
- Achievement tests should measure an adequate sample of the learning outcomes and subject matter content included in instruction.
- Achievement tests should be designed to fit the specific purposes for which the results will be used.
- Achievement test should be made as reliable as possible and should then be interpreted with caution.

The content of tests that are based on course objectives has a variety of benefits. The first is that it forces course designers to be clear about their goals. The second is that it allows test results to reflect how far students have progressed toward the learning outcomes. As a result, the course designer or teacher should create a syllabus based on the instructional objectives and choose books and materials that support those objectives.

Achievement tests are useful as they provide accurate information about students' learning and help teachers in making judgments about syllabus adjustments (Childs, 1989). However, doing so necessitates first defining and clarifying the instructional objectives. Teachers can then use the test results to learn about their students' abilities, needs, and achievement of the course objectives (Weir, 1995). Achievement tests are a crucial component of curriculum improvement since they help determine which objectives have been reached and where changes may be needed (Heaton, 1988).

Whatever the type of the test is, testing is central in any language program, because from the placement of the students in the program through their graduation, all kinds of program-related decisions are made using the tests as a basis. Testing directly affects teaching, because teachers use the test results to refine their course objectives, use of the materials and the activities used in the class, and if they are the testers, to prepare better tests (Brown, 1995).

2.4. Teacher-made test

According to Suharsimi Arikunto (2010), a teacher-made-test is a test that is written and created by a teacher in the classroom, so the test's validity and reliability are not the same as standardized tests. The effectiveness of this type of test is determined by the teacher's expertise and ability in creating it. This test is based on materials and specific objectives established by the teacher for his own students. Admittedly, teachers rarely evaluate and revise previously tested test items, they are unaware of the level of validity and reliability.

An Exploration into Teachers' Perceptions in Designing English Achievement Tests at Primary Schools in Thai Nguyen City, Vietnam

Brown (2003) stated that the teacher-made test can only be used in some of the teacher's classes. The advantage of using this type of test is that students are familiar with the assignment assigned by the teacher, allowing them to achieve higher scores than they would on a standardized test. Meanwhile, limited sampling, low reliability, subjective, low validity, high skill required, monotonous, and time consuming to prepare are some of the drawbacks of teacher-made tests.

III. METHODOLOGY

1. Participants

The participants of the study were 18 teachers who are currently working at primary schools in Thai Nguyen city. Among these teachers, there were two males and sixteen females. Due to this distribution, gender was not examined as a variable. Their experience as teachers' ranges between one and more than 10 years.

Three of these volunteers have Master's degrees in English language; three are studying for a Master's degree in English language at the School of Foreign Languages, Thai Nguyen University. Furthermore, all of the teachers had enough experience and passion in English language teaching to be fully qualified take part in this study.

2. Instruments and data collection

This study employed the quantitative method. More specifically, a set of questionnaire items were used to gather the teachers' perceptions in designing English achievement tests. By doing that, the researcher could get the descriptive data which are more reliable to answer the questions of the research.

In order to answer the research question, the questionnaire survey was conducted 1 month after the English language lecturers finished their English teaching course. 18 questionnaires for English language lecturers were distributed and 18 completed questionnaires were returned.

The purpose of the questionnaire in this study is to seek information about teachers' perceptions on the need of designing English achievement test and what the teacher thought about the representation of language skills and knowledge on their tests. More specifically, the questionnaire consists of 2 parts: Part 1 has 8 questions which aim to ask for the teachers' background information such as full name, gender, age, academic qualification, years of teaching English, English skills that they teach at the school, whether they conduct the test by themselves or from outside and training course attendance on language testing. Part 2 has 16 questions which utilize closed items in which teachers could express their perceptions about the English achievement tests they designed in their classroom. These questions are mostly based on a theoretical argument about the guidelines of English achievement tests.

3. Data analysis

The data obtained from 18 English lecturers were coded in to SPSS, then was analyzed separately in three steps. Firstly, the demographic information was computed to see the status of the respondents background. Secondly, the Cronbach's alpha values were computed to check the reliability of the questionnaire. Finally, the responses for each item in each part in the questionnaire were analyzed quantitatively.

IV. RESULTS

1. Demographics of teachers

From questions 1 to 8, the teachers' background information is sought out, including their full name, gender, age, academic qualification, year of teaching English, English skills taught at the school, whether they conduct the test internally or externally, and attendance at language testing training courses.

2. Teachers' perceptions in designing EAT

This section examines the perceptions that teachers use to drive their tests of students' achievement in their classes.

In order to establish teachers' perceptions in designing the achievement tests, descriptive statistics analysis was conducted by computing the mean scores for each item in four stages: planning, implementation, monitoring, and recording and dissemination in the teachers' questionnaire from question 1 to question 16. Table 1 presents the interpretation of the mean scores:

Table 1. The interpretation of the mean scores

Mean	Option	Degree
4.5-5.0	Strong agreement	Very high
3.5- 4.4	Agreement	High
2.5-3.4	Neutral	Moderate
1.5-2.4	Disagreement	Low
1.0-1.4	Strong disagreement	Very Low

Note: VH=Very high, H=High, M=Moderate, L=Low, VL=Very Low (Kucuk, 2007, p. 65)

An Exploration into Teachers' Perceptions in Designing English Achievement Tests at Primary Schools in Thai Nguyen City, Vietnam

The criteria for the mean scores adopted from Kucuk (2007) are shown in Table 1. Five Likert-scale criteria were used to assess the degree in which the respondents agreed with their principle in designing achievement tests. More specifically, the strongest agreement is between 4.5 and 5.0, followed by 3.5 to 4.4, and the undecided option is between 2.5 and 3.4. Finally, the strongest disagreement ranges from 1.0 to 1.4, while the disagreement ranges from 1.5 to 2.4.

In a nutshell, mean scores in Likert-scale criteria are used to assess a participant's attitude by determining to what extent they agree or disagree with a specific question or statement.

Stage 1: Planning

Table 2 reveals some teachers' perceptions that they use as a reference when preparing test of students' achievement in the classroom.

Table 2: Teachers' perceptions in the planning stage

Items	Propositions	Mean	SD	Degree
1	Teachers should first identify the purpose of the test and tests must be related to what students do in real class time when they design the test.	4.33	.485	H
2	Teachers should consider the attainment targets which the curriculum requests when they design the test.	4.33	.485	H
3	Teachers should use test specifications when they carry out the test.	4.22	.548	H
4	The appropriateness of test (tasks) should be checked by calling for peer comment or with reference to published guidelines (if these are available).	3.78	1.114	H
5	Teachers should give the students advance notice, so that the students will be able to prepare for the test.	4.28	.461	H
Overall mean score		4.188	.618	H

Note: VH=Very high, H=High, M=Moderate, L=Low, VL=Very Low

Propositions 1 and 2 receive the most positive comments in the first stage (Table 2), with the highest mean score ($M = 4.33$, $SD = .485$). Put it simpler, teachers agreed that they should first identify the purpose of the test, which is linked to what students do in practical class time when they design the test, and then they should identify the achievement targets that the curriculum requires when designing the test. Meanwhile, the least positive comment is for proposition 4 ($M = 3.28$, $SD = 1.114$), in which they consider that the appropriateness of the test should be reviewed by calling for peer comment or with reference to published guidelines. The rest ranges from 4.22 to 4.28, indicating the fact that teachers should use test specifications when they carry out the test and that teachers should give the students advance notice, so that they will be able to prepare for the test.

The overall mean score is accounted for at 4.11, demonstrating that all of the propositions submitted in the questionnaire were judged by the teachers to be part of their own evaluation principles. In descending order of means, the teachers who replied usually regarded the propositions given at this stage as part of their own test principles. Propositions 1, 2, 3, and 5 appear to be the principles that most teachers believe in and are most likely to be implemented in achievement tests. Thus, these findings suggest that teachers recognize the importance of developing a test strategy and understanding how to implement it.

Stage 2: Implementation

Table 3 presents teachers' perceptions in the implementation stage. It can be seen that the overall mean score of the implementation stage is ($M = 4.32$, $SD = .505$), showing that all of the propositions presented in the questionnaire were perceived by the teachers to be part of their own test principles.

An Exploration into Teachers' Perceptions in Designing English Achievement Tests at Primary Schools in Thai Nguyen City, Vietnam

Table 3: Teachers' perceptions in the implementation stage

Items	Propositions	Mean	SD	Degree
6	Teachers should inform the students of the reasons why they are being tested.	4.28	.575	H
7	Teachers should explicitly instruct the students how to do the test (tasks).	4.33	.485	H
8	Teachers should give students immediate feedback after they complete each test (task).	4.39	.502	H
9	Test (tasks) processes should be completed within a manageable time considering the given context.	4.28	.461	H
Overall mean score		4.32	.505	H

Note: VH=Very high, H=High, M=Moderate, L=Low, VL=Very Low

Table 3 depicts proposition 8 ranks first with the highest mean score ($M = 4.39$, $SD = 0.502$), followed by P 7 with the mean score of 4.33 ($M=4.33$, $SD=.485$) while the lowest mean score ($M= 4.28$, $SD= .575$) is accounted for the propositions 6 and proposition 9. This implies that teachers strongly agreed that they should explicitly instruct students on how to complete the test (tasks) and provide students with immediate feedback after completing each test (task). Additionally, the test (tasks) processes should be completed within a manageable time frame offered in the given context. In other words, propositions 8 and 7 reflect the principles that received the most support from teachers; these are the most likely to be applied to achievement tests. As a result, these findings indicate that teachers properly understand what to do and how to do it during the implementation stage of achievement tests.

Stage 3: Monitoring

Table 4 describes the perceptions that teachers use to monitor the implementation of their tests; this is essential to what they trust in the propositions provided below.

Table 4: Teachers' perceptions in the monitoring stage

Items	Propositions	Mean	SD	Degree
10	Teachers should construct a marking system as a part of the whole test process.	4.17	.618	H
11	Marking criteria should be connected with the aims of the test and the learner's characteristics in a given context.	4.33	.594	H
12	Teacher should use the results of test for revising their test	4.33	.485	H
13	Teachers should share the findings of assessment with other teachers.	4.17	.514	H
Overall mean score		4.25	.552	H

Note: VH=Very high, H=High, M=Moderate, L=Low, VL=Very Low

Table 4 shows that the mean scores for all items are above 4.1, the mean degrees are high, signifying the teachers' high agreement on the monitoring stage, and it could be argued that it more closely reflects the preferred principles of the teachers. In specific, propositions 11 and 12 achieve the highest mean score ($M =4.33$, $SD =.594$ and $SD=.485$), whereas the lowest mean score is accounted for proposition 10 and proposition 13 ($M= 4.17$, $SD=.618$ and $SD=.514$). This indicates that marking criteria should be connected with the aims of the test and the learner's characteristics in a given context, and teachers should use the results of the test for revising their test. In contrast, proposition 13 has the lowest mean score ($M = 4.17$, $SD=.514$), in which teachers agreed that they should share the findings of the test with other teachers.

An Exploration into Teachers' Perceptions in Designing English Achievement Tests at Primary Schools in Thai Nguyen City, Vietnam

Stage 4: Recording and Dissemination

Table 5 addresses teachers' perceptions when they record and share their students' achievement which has resulted from their implementation of the test.

Table 5: Teachers' perceptions in the Recording and Dissemination stage

Items	Propositions	Mean	SD	Degree
14	Schools should develop their own report system of students' progress and achievement.	4.22	.548	H
15	Teachers should be aware of their responsibilities for the output of their professional work.	4.06	.639	H
16	A formal review of a student's progress and achievement should be reported to the local education authority and the central government.	3.33	.970	M
Overall mean score		3.87	.719	H

Note: VH=Very high, H=High, M=Moderate, L=Low, VL=Very Low

Table 5 shows that the mean scores for the recording and dissemination stage are from moderate to high, ranging from 3.33 to 4.22. In specific, the most positive comment is for proposition 14 with the highest mean score ($M = 4.22$, $SD = .548$), whereas the lowest mean score ($M = 3.33$, $SD = .970$) is for proposition 16. It can be stated that claims 14 and 15 tend to reflect the favored principles of the teachers. In other words, teachers highly agree that schools should develop their own report system of students' progress and achievement, and teachers should be aware of their responsibilities for the output of their professional work. The exception is proposition 16, where the statistics indicate a nearly neutral or moderate stance towards the proposition; this could indicate that teachers are undecided about whether or not to fully support it.

V. DISCUSSION AND CONCLUSION

This section brings together all aspects of results from the study, answers the research questions and discusses the findings. The discussion is thus guided by the aims of the exploration into teachers' perceptions in designing English achievement tests at primary schools in Thai Nguyen city

Stage 1: Planning

When planning English achievement tests, the teachers seem to have a strong understanding of what is valid. This leads to a variety of problems that mirror features of their testing context. Concerning the purposes of tests (Propositions 1 and 2), it appears that teachers understand the importance of planning for achievement tests. They first organize their test by comparing the students' achievement tests to the curriculum's objectives, and then reflecting on the students' learning experiences. These findings suggest that the teachers' perceptions of the test are consistent with one of the achievement test's characteristics. The findings are in line with Arkoudis and O'Loughlin (2004), who claimed that teachers should deeply understand the process of testing and assessment of learning for academic purposes. It appeared that the teachers were primarily concerned with testing for "educational purposes".

Furthermore, the teachers truly believed that they need to apply their judgments based on a certain level of professional expertise because they were sometimes forced to choose between the published technical specifications and their own professional judgment (Davison, 2004, p. 316). This also demonstrates that, while the national curriculum has a significant influence on what happens in the classroom, it is nonetheless localized to some extent through the teachers' own perception. Thus, teachers in the middle of the five-staged cline of teachers could be classed as "assessors" (Davison, 2004, p. 325). This indicates that they intend to reconcile the required criteria with contextual considerations as principled yet pragmatic experts. Sharing the same view, several researchers (Davison, 2004, Shim, 2009) highly agree that teachers should assume a role as assessors who think critically, creatively, and logically in order to assess a student's performance in order to give the teacher an idea of how students learn, their attitudes, what skills they have, what skills they are working on, and how they are progressing.

As indicated in the preceding section (Section 2.1.3, Chapter II), for a test to be called "good", it must be accomplished through appropriate procedures; that is, it cannot be anticipated to come up with a real result or feedback on what the students know or can achieve unless the test is done in a rational way (Proposition 5). Teachers' opinions are typically in accordance with the concepts stated in this study during the implementation stages; they have a good understanding of what a legitimate method and ethical evaluation is.

An Exploration into Teachers' Perceptions in Designing English Achievement Tests at Primary Schools in Thai Nguyen City, Vietnam

Stage 2: Implementation

Teachers' perceptions are typically in line with the principles stated in this study during the implementation stages; they have a good understanding of what a valid procedure is. This raises a number of issues which should be addressed. Teachers considered that students should understand how to accomplish the tasks and attempted to finish the test within the time given. They appeared to be acting as an "interpreter" (Davison, 2004, p. 325) who followed the guide's suggestions. The results for specific guidance on how to complete the test tasks (Proposition 7) and task completion within an expected time (Proposition 9) reveal that the teachers' test principles are compatible with one of the features of EAT, namely that EAT is impacted by both the national and school curriculum.

Stage 3: Monitoring

During the monitoring stages, teachers have a firm understanding of how to use test findings and how to acquire consistent marks or information on students' achievement or performance. However, there are a number of difficulties that must be addressed. First, it shows that the teachers were confident in their use of test results. They stated that they utilized test findings largely to reflect on what they did in the classroom and then to improve their teaching (Proposition 12). As a consequence of these findings, the teachers' perceptions were found to be mostly similar to the argument that testing is not an isolated event but is linked to a full set of variables that interact in the educational process (Shohamy 1993; Shohamy et al 1997). Hughes (1993) defines this process as material development, syllabus design, changes in teaching methodologies, and learning and test-taking strategies. Teachers' understanding of the washback effect was restricted in this study because their responses indicated that they used the assessment findings primarily to revise their teaching approaches rather than to analyze other areas such as syllabus design and textbooks used. This might be due to the fact that they were given standardized syllabi and the MOET created most of their materials. This may have reduced teachers' understanding of the importance of challenging or revising them. Second, the teachers were concerned about the negative usage of test results. They thought that the test results could be used negatively in the classroom in very few, if any, instances. They were, however, unaware of the use of test results by others involved with EAT. This lack of awareness of the use and misuse of EAT results could be solved by improving their professional competence.

Stage 4: Recording and Dissemination

In recording and dissemination stage, the teachers believed that their school should develop their own report systems and that they should participate the development process (Propositions 14 and 15). Proposition 16 relates to the power relationships of tests. According to the findings of the reporting of students' progress and achievement, teachers were moderately worried about whether or not they should report their students' educational results to the local educational authority or the government.

In conclusion, the study's findings demonstrated that the majority of teachers have strong perceptions about EAT and have a good understanding of testing principles; therefore, they conducted their tests using an appropriate process while keeping in mind the English teaching and testing context. It can be noted that the findings from the study have shed light on important factors which relate to the effect of the process of test preparation and test construction. Thus, it is further hoped that this study will be a contribution to the ongoing efforts to provide more validity evidence for in-house English language tests. Particularly, the researcher hopes that this study will be useful not only for those currently involved in English teaching and testing in schools, such as teachers, principals, and students, but also for other stakeholders interested in English language testing, such as students' parents, supervisors, administrators, and policymakers. Furthermore, this research might serve as a springboard for anyone interested in conducting more extensive research in this field.

VII. STATEMENTS AND DECLARATIONS

The author declares that this research received no external funding and no conflict of interest

REFERENCES

- 1) Bachman, L. F. & Palmer, A. S. (1996). *Language Testing in Practice*. UK: Oxford University Press.
- 2) Basanta, C. P. (1995). Coming to Grips with Progress Testing: Some Guidelines for Its Design. *Forum*, 33, 55-64.
- 3) Brown, J. D. (1995). *The Elements of Language Curriculum*. Boston: Heinle & Heinle Publishers.
- 4) Brown, H. (2003). *Language Assessment Principle and Classroom Practice*. New York: Pearson Education, Inc.
- 5) Brown, J. D. (1996). *Testing in Language Programs*. New York: Prentice Hall Regents.
- 6) Childs, R. A. (1989). *Constructing Classroom Achievement Tests*. Được truy lục từ ERIC Digest: <http://www.ed.gov/databases/ERIC Diuests/ed 315426.html>
- 7) Cohen, R. J. (1988). *Psychological Testing and Assessment*. California: Mountain View.
- 8) Cronbach, L. (1971). *Educational Measurement*. Washington, DC: American Council on Education.
- 9) David, A. (1990). *Principles of Language Testing*. Padstow: Cornwall: Basil Blackwell Ltd.

An Exploration into Teachers' Perceptions in Designing English Achievement Tests at Primary Schools in Thai Nguyen City, Vietnam

- 10) Fernandes. (1984). *Testing and Measurement*. Jakarta: Nasional Education Planning.
- 11) Finnachiaro, M., & Salo, S. (1993). *Foreign Language Testing*. New York: Regents Publishing Co.
- 12) Glenn, F. (2010). *Practical Language Testing*. Britain: Hodder Education, An Hachette UK Company.
- 13) Hawkey, R. (2013). "Consequential validity" in *Examining listening: Research and practice in assessing second language listening*. Vol. 35. eds. A. Geranpayeh and L. Taylor (Cambridge: Cambridge University Press), 273–302.
- 14) Heaton, J. (1988). *Writing English Language Tests*. Hong Kong: Longman Group UK Ltd.
- 15) Heaton, J. B. (1990). *Classroom Testing*. New York: Longman Inc.
- 16) Hughes, A. (1989). *Testing for Language Teachers*. Cambridge: Cambridge University.
- 17) Hughes, A. (1995). *Testing for Language Teachers*. Cambridge: Cambridge University Press.
- 18) Julia, Brannen.(1992). *Mixing Methods: qualitative and quantitative research*. London: Taylor & Francis.
- 19) Penny, U. (1996). *A Course in Language Teaching: Practice and Theory*. Cambridge: Cambridge University Press.
- 20) Fernandes, H, J, X (1984). *Testing and Measurement*. Jakarta: Nasional Education Planning.
- 21) Field, J. E. (2013). "Cognitive validity" in *Examining listening: Research and practice in assessing second language*. eds. A. Geranpayeh and L. Taylor (Cambridge: Cambridge University Press), 77–151.
- 22) Lim, G. S., and Khalifa, H. (2013). "Criterion-related validity" in *Examining listening: Research and practice in assessing second language listening*. Vol. 35. eds. A. Geranpayeh and L. Taylor (Cambridge: Cambridge University Press), 303–321.
- 23) Linn, R. L., & Gronlund , N. E. (1995). *Measurement and Assessment in Teaching*, Seven edition. London: Merrill.
- 24) Sardar, S, "Test, Measurement, and Evaluation"[Online]. <https://idoc.pub/documents/dr-sanjit-sardar-test-measurement-and-evaluationpdf> accessed on 25/4/2020.
- 25) Taylor, L., and Geranpayeh, A. (2011). Assessing listening for academic purposes: defining and operationalizing the test construct. *J. Engl. Acad. Purp.* 10, 89–101. doi: 10.1016/j.jeap.2011.03.002
- 26) Van, H. V. (2007). The current situation and issues of the teaching of English in Vietnam. *Teaching English to speakers of other languages in the internationalization of higher education in Vietnam*, 3(4), 1-12.
- 27) Weir, C. (2005). *Language testing and validation: An evidence-based approach*. Basingstoke: Palgrave Macmillan.



There is an Open Access article, distributed under the term of the Creative Commons Attribution – Non Commercial 4.0 International (CC BY-NC 4.0) (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits remixing, adapting and building upon the work for non-commercial use, provided the original work is properly cited.