

ADAMA SCIENCE AND TECHNOLOGY UNIVERSITY
SCHOOL OF HUMANITIES AND LAW
DEPARTEMENT OF ENGLISH
(POST-GRADUATE
PROGRAM)

A STUDY ON TEACHERS' PRACTICE OF DESIGNING
ENGLISH LANGUAGE TESTS AT CLASS ROOM LEVEL:
SHIRKA GOVERNMENT HIGH SCHOOL IN FOCUS.

BY
LEGESSE REGASSA

MAY, 2015

A STUDY ON TEACHERS' PRACTICE OF DESIGNING
ENGLISH LANGUAGE TESTS AT CLASS ROOM LEVEL:
SHIRKA GOVERNMENT HIGH SCHOOL IN FOCUS.

A THESIS SUBMITTED TO THE SCHOOL OF
HUMANITIES AND LAW DEPARTMENT OF ENGLISH IN
PARTIAL FULLFILMENT OF THE REQUIREMENTS FOR
THE DEGREE OF MASTERS OF ARTS IN TEACHING
ENGLISH AS A FOREIGN LANGUAGE (TEFL).

BY: LEGESSE REGASSA
TO: JEYLAN AMAN (PHD)

MAY, 2015

APPROVED BY THE BOARD OF EXAMINERS:

ADVISOR

SIGNATURE

DATE

EXAMINER

SIGNATURE

DATE

EXAMINER

SIGNATURE

DATE

STATEMENT OF THE AUTHOR

First, I declare that this thesis is my original work and that all sources of materials used for this thesis have been heartily acknowledged. This thesis has been submitted in partial fulfillment of the requirement for MA degree in teaching English as foreign language (TEFL) to school of humanities and law, Department of English at Adama Science and Technology University.

Name: Legesse Regassa

Place: Adama Science and Technology University

Date of submission: 11/5/2015

Signature: -----

ACKNOWLEDGMENTS

First of all, I would like to thank my God who always helps me in all circumstances. Next, my special gratitude goes to my M.A. thesis advisor Dr. Jeylan Aman, for his unreserved guidance, material support and constructive comments throughout the course of this study. My heartfelt gratitude also goes to my wife teacher Zebideru Dinku, for her advice and support in writing the manuscript on computer. Then, my gratitude goes to my friends who are teaching English at Shirka High School for providing me with the necessary information and sharing their experiences when I was collecting data for my research work. Finally, I would like to thank my family for their financial support. And to those I have not mentioned their names here, I say many thanks.

TABLE OF CONTENTS

CONTENTS PAGE

Acknowledgments-----	i
Table of Contents-----	ii
List of Tables-----	vi
Key words-----	vii
Abstract-----	viii

CHAPTER ONE

Introduction-----	1
1.1. Back ground of the study-----	1
1.2 Statement of the problem-----	3
1.3. Objective of the study-----	4
1.3.1 General objective -----	4
1.3.2 Specific objectives-----	4
1.4 Research Questions-----	4
1.5 Significance of the study-----	5
1.6. Delimitation (Scope) of the study-----	5
1.7.Limitation of the study-----	5
1.8. Organization of the paper-----	6

CHAPTER TWO

Review of Related Literature-----	7
2.1. Introduction-----	7
2.2. International Related Research Works Conducted on Language Testing Practice -----	8
2.3. Local Research Works Conducted on Language Testing Practice-----	10
2.4. Effective Test Item Construction Techniques-----	10
2.5. Items Commonly Used for Tests of Achievements-----	14
2.5.1. Structured Response/Select Items-----	14
2.5.2. Constructing Response/supply Items-----	19
2.5.3. Selecting Passages for Test Purpose-----	20

2.5.3.1. Guide Lines for Selecting of Specific Kinds of Passages-----	20
2.6. Principles for designing Effective Class-room Test-----	21
2.7. Practical Steps for Test Construction-----	23
2.8. Conclusion-----	25

CHAPTER THREE

Research Methodology -----	26
3.1. Research Method-----	26
3.2. Research Site-----	26
3.3. Population and Sampling -----	27
3.4. Instrument of data collection-----	27
3.4.1. Document analysis-----	28
3.4.2. Teachers' interview-----	28
3.5. Procedures of data collection -----	28
3.6. Procedures of data analysis-----	29

CHAPTER FOUR

Data Analysis and Discussion-----	31
4.1. Analysis and Discussion of Test Items-----	31
4.1.1. Analysis and Discussion of multiple choice items-----	32
4.1.2. Analysis and Discussion of Teachers' Techniques in Selecting Reading Passages-----	36
4.2. Analysis and Discussion of Teachers' interview Responses on each question-----	38
4.2.1. Analysis and discussion of teachers' responses regarding strategies and techniques they use in designing classroom tests-----	38
4.2.2. Analysis and discussion of teachers' responses regarding principles that they follow in designing English language tests at classroom level-----	40
4.2.3. Analysis and discussion of teachers' responses regarding type of test items they frequently use in designing class room tests-----	41
4.2.4. Analysis and discussion of teachers' responses regarding factors that affect their practice in designing effective tests at classroom level-----	41
4.3. Discussion of Teachers' overall Interview Responses-----	41

CHAPTER FIVE

Summary, Conclusions and Recommendations-----	45
5.1. Summary-----	45
5.2. Conclusions-----	48
5.3. Recommendations-----	49
References-----	50
Appendix A: Checklists for Analyzing Teacher-Made Test Items-----	53
Appendix B: Semi-structured Interview items for Teachers-----	58
Appendix C: Sample of teachers-made exams-----	60
Appendix D: Interviewees responses on each interview questions-----	61

LIST OF TABLES

Table 1: Analysis of Multiple Choice Items

Table 2: Analysis of Teachers' Techniques in Selecting Reading Passages

KEY WORDS/PHRASES/

The following key words or phrases were used interchangeably in the study:

- **Class room level** used as **department level**.
- **Teacher-madetest** used as **class room tests**.

ABSTRACT

The purpose of this study was to assess teachers' practice of designing English language tests at classroom level in Shirka high school. Mixed research method was used to carry out this study. The subjects of the study were ten grade tenth English language teachers. The samples were selected using comprehensive sampling technique. The data gathering tools were document analysis and semi-structured interview. The findings of the study revealed that the teachers have no common techniques that they use in designing exam items at department level, they do not use any written exam plan personally and at department level before constructing exam items and did not use proper principles to be employed when they design classroom test. The findings of the study further revealed that the teachers use multiple choices item types repeatedly than the other types of test items in their exam and almost all of the teachers did not take any training regarding test designing practice at district education offices, regional education offices and ministry of education level to update their experiences of test designing skills. Finally, recommendations were forwarded based on the major findings of the study to the teachers and concerned body to bring necessary changes on teachers test designing practice.

Chapter One

Introduction

1.1 Background of the Study

According to Hughes (1989) explanation tests are very useful, especially in language teaching. They are the means to show teachers how much the learners have learnt during the teaching learning process. This can be possible if and only if tests are designed by fulfilling the quality of good tests based on their purpose. For tests to cover essential qualities, it is the duty of test designers to be acquainted with good test designing experiences at different levels. Though tests are constructed at different levels by different experts, teachers are one of those test developers to take responsibility of making decision about test designing at classroom levels.

Teachers make decisions about test designing practices on a day-to-day basis. Their decisions are influenced at large extent by their past testing and assessment experiences. The experiences they get from past and present work can have negative or positive results on their testing and teaching practices. For example, learning from past mistake helps teachers to change their attitude of past experiences. Thus teachers can change their mind to what they have observed, make a decision, and carry out appropriate actions on their work.

Therefore, because testing is one of the measurement tools that help teachers to check their failure and success in teaching process, teachers need to improve their practice different from their own past weak testing experiences and continue their current and future assessment practices in a better way. There are a lot of problematic factors that can underneath of tests constructed and implemented at classroom level. For example absence of stationery in schools, too little time given for test takers, and teachers' practice of designing quality tests are some problems experienced in test designing activities.

Though many surveys of language testing practice have been conducted, very few of them, are focusing on the test developers' practices. Thus, it is better to deal with two retrieved studies from the language testing literature, examining test developer's practices, conducted by Alderson and Buck (1993) and Alderson (2010). In one such study, (Alderson and Buck 1993) examined the testing practices of the English as second language examination boards in the United Kingdom context. The purpose of their study was to examine what standards were

maintained by the English as second language examination boards in the United Kingdom, to what extent all English as second language examining authorities followed the same or similar procedures, and whether procedures could be improved. Their study was divided into two phases. In the first phase, a letter was sent to the English as second language examination boards, asking three broad questions in relation to the standards to which English as second language examination boards followed, and the procedures that they followed to estimate test reliability and to ensure test validity.

During the second phase of the study, the two researchers used a structured questionnaire to investigate the testing practices of twelve examination boards, touching upon such areas as test syllabus, test construction, validation, administration, marking, reports and test revision. Their study concluded that there was no common agreed set of standards that these English as second language examination boards maintained, and different examination boards did different things, with different degrees of situation, to monitor the quality of their examinations. The two researchers therefore further concluded that time was ripe for United Kingdom examination boards and United Kingdom language testers to develop a set of standards all English as second language tests should follow.

In the other study, Alderson (2010) surveyed the testing practices of the examination boards or agents responsible for developing and delivering aviation English language tests. Similar to the 1993 study, this study was also divided into two stages. During the first stage, a 'filter' questionnaire was employed, asking broad questions such as respondents' understanding of the International Civil Aviation Organization scales and a brief section on quality control procedures.

During the second stage of the study, a structured follow-up questionnaire was developed on the basis of the guidelines for good practice in language testing and assessment aiming to survey in detail, the testing practices of the test developers. The areas in testing practice under survey included test purpose and specifications, test design and item writing, rating procedures, test administration, and test review. The respondents in this study were twenty-two persons or organizations from all over the world that were responsible for developing or delivering aviation English language tests. The study identified worrying problems for aviation English test

development and validation, and proposed that standards such as the language testing and assessment guidelines were adjusted to suit the aviation context.

These two studies are both significant in the sense that they portrayed the testing practices of the examination boards in a certain context. Furthermore, the 1993 study, as the first systematic study of such a nature that appeared in language testing, helped to inspire many test developers following discussions of standards and professionalism in the language testing community. In contrast, the 2010 study directed their attention to the quality and professionalism in developing aviation English test, a genre of very high-stakes language test. However, neither of these two studies addressed such issues like teachers' practice of designing English as foreign language tests at classroom levels, test fairness, and test use in the research design. All are becoming very important issues nowadays in language testing. Furthermore, neither of these two studies investigated the reasons or rationales behind the inadequacies of test developers.

Even though different problems can be related to tests designing activities, teachers' practice in designing quality tests at classroom level is one of the big problematic areas in tests construction process. Yet it seems that only few researchers have made this area the focus of their studies. For instance, Alderson (2011) revealed that though many surveys of language testing studies have been conducted, very few of them are focusing on the teachers' practice of designing tests probably because of data inaccessibility or lack of scientific rigor in such studies. Cheng (2005)' pointed to the urgent need for teachers to understand the nature of testing, to refine their own classroom assessment without unquestioningly accepting or following large-scale testing. According to Brindley (2007), studies of classroom assessment practices used by teachers of English as foreign language have been relatively limited compared with the fairly large body of studies of large-scale language testing practices in the field.

1.2. Statement of the Problem.

Teachers' practice of designing tests affects their classroom test quality which in turn affects students' learning. According to Rodriguez (2004), the teacher-made classroom test (exam) is thought to have strong influence on student achievement, study patterns, self-perceptions, attitudes, efforts, and motivation to learn. If the exam is designed and implemented in the right way, it can encourage students toward their educations, but if the exam is not designed in the

way it is expected, it can affect students' achievements. Therefore, without solving the problem related with test designing practices, the goal to bring students to effective test achievement is unthinkable.

On other hand, regarding the importance of classroom test and cautions to be taken when designing it, Axman(1989) stated that achievement tests should be well suited to provide teachers with objective feed back as to how much students are learning and understanding. Expanding his explanation he says, classroom achievement tests, if used carefully can provide not only the evaluation of the knowledge levels of individual students, but also information about how those students are compared with each other in the classroom. They seldom provide feedback on the mastery or non-mastery of the full range of specific skills taught in a given classroom.

Though some research works have been conducted internationally on similar problems, there are inadequate research works that have been conducted locally on this area. For instance, Kifle (1995) and Argaw (2004) have conducted studies on the National Examination of grade 10. The specific objectives of these studies were to see the content validity of the examinations. The results of the studies showed that the examination lacked content validity. Yeshambel (2011) and Samuel (2011) have also assessed the content validity of grade 8 and 9 English language examinations. They concluded that both examinations lacked content validity.

Even though, these local research studies were conducted on English language examinations at different grades, they did not have direct connection with teachers' test designing practice. As it can be understood from this fact, it seems that no many research works have been conducted on teachers' practice of designing English language tests at classroom level at local level. This reveals that it is the right time to carry out research on this title. Therefore, this study attempted to examine teachers' problems of designing classroom tests at Shirka government high school.

1.3 Objectives of the Study

1.3.1 General Objective

The main purpose of this study was to examine English language teachers' practice of designing English language tests at classroom level.

1.3.2 Specific Objectives

The study was specifically aimed to:

1. Explore strategies and techniques English language teachers use in designing classroom tests;
2. Examine the principles that English language teachers follow in designing English language tests at classroom level.
3. Identify the types of test items that are most used by English language teachers;
4. Examine factors that affect teachers' practice in designing effective tests at classroom level.

1.4. Research Questions

The study attempted to answer the following research questions.

1. What are the strategies and techniques English language teachers use in designing classroom tests?
2. What are the principles that English language teachers follow in designing English language tests at classroom level?
3. Which type of test items are most frequently used by English language teachers?
4. What are the factors that affect teachers' practice in designing effective tests at classroom levels?

1.5 Significance of the Study

Investigation of teachers' practice of designing English language tests has its own benefits for some pedagogical and research purposes. Pedagogically, the findings that will be obtained from this study are hoped to contribute some benefits for the English language teachers teach at high school level. The findings help these English language teachers to review their skill of designing tests at classroom levels carefully as pedagogical purpose. As a result, they will be motivated to

study and use different strategies, principles, guidelines that help them in designing English tests at classroom level for students.

On the other hand, the importance of this study for research purpose is that it is hoped to provide helpful insights to other researchers who want to carry out further studies on the area about teachers' practice of designing classroom language tests. The clues that reveal some facts about teachers' problem in designing tests show directions to other researchers who want to conduct studies on similar cases.

1.6. Delimitation of the Study

According to Blaxter (2006) conducting research in researcher's work place has some advantages for his study successfulness. These advantages are ease of access to the data, be able to do research in work time, receive financial and other support from his employer, his insider knowledge about the topic and colleagues may provide him with lots of useful information. Taking the scholar's suggestion into consideration, the study was delimited to Shirka High School, Shirka district, Arsi zone, Oromia Regional state where the researcher is working at the moment.

Moreover, the rationale behind delimiting the scope of the study under the topic "The study on teachers' practice of designing English language tests at class room level" were the seriousness of the problem that affect many English language teachers' test quality, the newness of the topic; that is the topic has not been researched before in this area as the researcher's knowledge concerned, researcher's capital status to afford for the study, applicability of the results that depends on the researcher capability and availability of resources to implement were taken into consideration to delimit the scope of the study under this topic.

1.7. Limitation of the study

Ten teachers who have been teaching English in their regular classes were selected for sample size from one high school. The sample size for this study could not warrant generalization to other English language teachers in other schools. This is because other teachers might not exercise this type of testing practice in their school. In fact the comprehensiveness of the study would have been increased if it had included more schools as study sites, but constraints of time and money

forced the researcher to select only one high school. On top of this, literature sources which directly related to the topic under study were not adequately available in the internet. This situation threatened the researcher to see the research gap existed deeply in the study area, especially studies conducted at local level. As a result, the study did not clearly show the gap existed in the study area. On the other hand, the researcher believe that some of the interviewees did not list all their poor practice of testing as they were seen trying to conceal some of their weakside when interview was made with them in their school. This situation could affect the data gathered and put a little bit limitation on the reliability of the study.

1.8. Organization of the paper

This study paper was organized to have five chapters. The first chapter dealt with introduction of the study, which contains back ground of the study, statement of the problem, objectives of the study, significance of the study and scope of the study. Chapter Two discussed the review of related literature. Chapter three was concerned with research methodology, specifically, the research design, the subjects of the study, sampling technique, data gathering instruments. Chapter four discussed about methods of data analysis and interpretation. Finally chapter five presented summary, conclusions and recommendation of the study.

Chapter Two

Review of Related Literature

The primary purpose of this chapter is to review the literature on teachers' classroom test designing practices. To do so, techniques, principles and strategies help them in designing better classroom tests for their students are presented in short and clear ways. The chapter begins with an introduction part that deals with the importance of review of literature in this study. Then, it presents important literatures that can provide valuable information for English language teachers on designing and writing classroom tests effectively for their students. Finally, the chapter presents the conclusion that shows the reader the overall concepts discussed in chapter.

2.1 Introduction

Cheng, L. (2005) said that teachers have primary responsibility for designing classroom tests and evaluate the weakness and strength of their own instruction and evaluate the learning progress of their students. Teachers' decision can directly influence student achievement, study patterns, self-perceptions, attitudes, effort, and motivation to learn. According to Rodriguez (2004) ineffective test designing practices can occur without clarifying the problems between those practices and the test designing competencies of classroom teachers. It is good testing that can resolve such problems. Testing is one of the teacher's most important tasks. What teachers test, how they test and why they use test; show students about what is worth learning, how it should be learned, and how well they are expected to learn it. As a result, teachers need to increase working better to bring their classroom testing practice with a continuum of benchmarks and standards, and students are expected to study in a way that they can be benefited from their studying. Clearly, classroom tests can shape much of what is happening in classrooms (Horn, & Ramos, 2000).

Teachers design classroom tests for a variety of purposes. Some bring students a sense of success and fairness, while others strengthen student perceptions of their strength and weakness. Regardless of their intended purpose, classroom tests directly or indirectly influence students' future learning, achievement, and motivation to learn. Testing practice has been discussed in language testing literature, and has been approached from different perspectives by language testing researchers. A common approach to addressing this issue, for example, is to discuss how a language test should be developed, administered, and evaluated. Therefore, this discussion

primarily focuses on good practice in each and every step in the testing cycle. It includes, for instance; test specifications, item writing, test administration, marking, and reporting test results. Though the idea in this paper concentrates on discussing good practice in designing English language tests, it also attempts to discuss some assumptions and suggestion forwarded by different English as second language researchers upon test developing practices and approaches.

2.2. Effective Test Item Construction Techniques

As Madsen, H. (1983) education is a process that helps the students change in many ways. Some aspects of the change are intentional whereas others are quite unintentional. Keeping in view this assumption, one of the important tasks of language teachers is to decide, as far as possible how they want their students to be changed and to determine their role in this process of change. Secondly, upon the completion of a particular unit/ course, the teachers need to determine whether their students have changed in the desired manner. So to determine students change they need to design language tests following effective test item construction techniques. Some of these techniques are the following ones. According to Hughes, A. (1989) language teachers need to follow the following techniques when they design tests.

2.2.1 Planning the Test

Test planning includes several activities that need to be carried out by the teacher to devise a new test. As a first step, the teacher must draw up a test “blue print”, specifying: the content and objectives of the test; types of items; practice exercises; time limit etc.

2.2.1.1. Objectives of the test

The basic objective of an educational achievement test is to assess the desired changes brought about by the teaching-learning process. Obviously, each subject demands a different set of instructional objectives. For example, major objectives of the subjects like sciences, social sciences, and mathematics are: knowledge, understanding, application and skill. On the other hand the major objectives of a language course are: knowledge, comprehension and expression.

Knowledge objective is considered to be the lowest level of learning whereas understanding, application of knowledge in sciences / behavior sciences is considered higher levels of learning.

As the basic objectives of education are concerned with the modification of human behavior, the teacher must determine measurable cognitive outcomes of instruction at the beginning of the

course. The evaluation process determines the extent to which the objectives have been attained, both for the individual students and for the class as a whole. Such an evaluation provides feedback that can suggest modification of either the objectives or the instruction or both. Some objectives are stated as broad, general, long-range goals, e.g., ability to exercise the mental functions of reasoning, imagination, critical appreciation. These educational objectives are too general to be measured by classroom tests and need to be operationally defined by the class room teacher (Bachman, F. and Palmer, A. 1996).

2.2.1.2 The Content:

The outline of a classroom test should include details of test content in the specific course. Moreover, each content area should be weighted roughly in proportion to its judged importance. Usually, the weights are assigned according to the relative emphasis placed upon each topic in the textbook. The median number of pages on a given topic in the prescribed books is usually considered as an index of its importance.

2.2.1.3 Judgments of Experts

Bachman, L. (1990) suggests that to devise a classroom tests, the advice and assistance of subject-matter experts, serving as consultants, can prove to be of immense importance. The best way to seek consultants' judgments is to submit to them a tentative outline prepared after studying the instructional objectives as stated in representative curriculums and of the subject-matter content as indicated by up-to-date and widely used textbooks.

2.2.1.4 Preparing the Test Blue Print

According to Kopriva, R.J. (2008) the blue print is meant to ensure content validity of the test. It is the most important characteristic of an achievement test devised to determine the **GPA** at the end of a unit/ term or course of instruction. The test may be based on several lessons or chapters in a textbook, reflecting a balance between content areas and learning objectives. The test blueprint must specify both the content and process objectives in proportion to their relative importance and emphasis in the curriculum. Depending on the purpose of the test and the instructional objectives, the test may vary in length, difficulty, and format (objective, essay, short-answer, open-book, or take-home).

2.2.1.5. Preparing the Test According to Plan

The next step after planning the test is preparing it in accordance with the plan. This step mainly deals with development of items and organizing them in the form of a test.

2.2.1.6. Test Administration and Use

All pupils must be given a fair chance to demonstrate their achievement. Physical and psychological environment be conducive to their best efforts. Control all factors that might interfere with valid measurement: Adequate workspace, quiet, proper light and ventilation are important. Pupils must be put at ease, tension and anxiety should be reduced to the minimum.

2.2.1.7. Scoring the test

Lyman, H.B. (1998) states that if the pupils' answers are recorded on the test paper, the teacher may make a scoring key by marking the correct answers on a blank copy of the test. When separate answer sheets are used, a scoring stencil is a blank answer sheet with holes punched where correct answers should appear. Before scoring procedure is used, each test paper should also be scanned to make sure that only one answer was marked for each item. Any item containing more than one answer should be eliminated from scoring. In scoring objective tests, each correct answer is usually counted as one point. When pupils are told to answer every item on the test, a pupil's score is simply the number of items answered correctly. Short answer questions may sometime require awarding partial credit and may pose some problem in scoring. However, a detailed key may be prepared in advance to avoid confusion. For each question and for the test as a whole, the examiner may make a tally for each kind of error that the examinees make. A summary of these errors could then be used to plan instructional activities.

2.2.1.8. Test Results

As Lyman, H.B. (1998) explains, the score obtained by a pupil has no meaning at all and cannot be directly interpreted. If a student obtains 75 marks out of 100, it tells us neither how s/he compares with other students nor what he knows nor what he does not know. The simplest form of meaning that teachers often provide to the test score is by assigning ranks to the scores. It however, becomes more interpretable when one knows the total number of the students in the class. Often grades are assigned to give meaning to the raw scores by comparing individual performance with the whole group that has taken the test. In educational institutions most often it

is the class. For criterion referenced test, of course, absolute grading is used (Lyman, H.B. (1998)).

2.2.1.9. Reviewing Test Results

It is important that the teacher returns and review test results with the students. Feedback on the performance has a special value for motivating students for improvement. Moreover, learning from one's mistake is usually very effective.

2.2.2. General Recommendations that Apply to All kinds of Test Developing Process

Kubiszyn, K., & Borich, G. (1984) suggest the following recommendation that can be applied to all kinds of test developing process. Keep the test plan in view as test exercises are written. Items should be addressed to the cells in the blueprint / the test plan. Draft the test items some time in advance, and then review them. Have test items examined and critiqued in, the light of the rules for writing items, by one or more colleagues. Check if the item has an answer that would be agreed upon by experts. If possible one of the experts may take the test and the responses of the expert are compared with the keyed answers. This way any error can be detected and the test developer may put confidence in the key thus finalized. Prepare a surplus of test exercises, so that an adequate number of good items will be available for the final version of the test. Assembling a test Items after having written and selected they are organized in the form of a test.

As (Fulcher, G. 2010) explanation, items of the same format may be placed together. Each item type requires specific set of directions and a somewhat different mental set on the part of the examinee. So far as possible, within item type, items dealing with the same content may be grouped together. The examinee will be able to concentrate on a single domain at a time rather than having to shift back and forth among areas of content. Furthermore, the examiner will have an easier job of analyzing the results, as it will be easier to see at a glance whether the errors are more frequent in one content area than the other. Items may be so arranged that difficulty progress from easy to hard. Items should be arranged in the test booklet so that answers follow no set pattern.

The directions should be simple but complete. They should indicate the purpose of the test, the time limits and the score value of each question. Write a set of directions for each item type that

is used on the test specifying what the respondent is expected to do and how one is required to record the response. Separate answer sheets, which are easier to score, can be used at high school level and beyond.

2.3. Items Commonly Used for Tests of Achievement

Two major types of items are commonly used for tests of achievement. These are:

1. Structured Response / Select items
2. Constructed Response / Supply items

2.3.1. Structured Response / Select Items

In the select type items, as the name suggests the examinee is required to select the correct answer from amongst the given or structured options. They are often called objective items.

They include: Alternate Response, Multiple-choice and Matching. Construction of test items is a crucial step for the validity of a classroom test is determined by the extent to which performance to be measured is called forth by the test items. It is not enough to have knowledge of subject matter, defined learning outcomes, or a psychological understanding of the students' mental states. The ability to construct high-quality test items requires knowledge of the principles and techniques of test construction and skill in their application. Objective test forms typically measure relatively simple learning outcome.

2.3.1.1 Alternative Response Items

Alternative response item, by definition is the one that offers two options to choose from. They often consist of a declarative statement that the examinee is asked to mark true or false, right or wrong, correct or incorrect, yes or no, agree or disagree, or the like. Incomplete sentences providing two options to choose from to fill in the blank also fall in this category. Very common use of such items is to test the knowledge of grammar. Appropriate use of 'tense' and also, contextual meaning of words or spelling mainly of words that sound alike. In each case there are only two possible answers. The most common form it takes is True-False questions.

2.3.1.1.1 True-False Items

According to Heaton, J.B. (1988) most common use of the true-false item is in measuring the examinee's ability to identify the correctness of statements of fact, definitions of terms, statements of principles, and the like, also to distinguish fact from opinion. True-false tests include numerous opinion statements to which the examinee is asked to respond true or false. There is no objective basis for determining whether a statement of opinion is true or false. In most situations, when a student is the respondent, he guesses what opinion the teacher holds and marks the answers accordingly. This, of course, is not desirable from all standpoints, testing, teaching, and learning. An alternative procedure is to attribute the opinion to some source, making it possible to mark the statements true or false with some objectivity. This would allow measuring knowledge concerning the beliefs that may be held by an individual or the values supported by an organization or institution. Another aspect of understanding that can be measured by the true-false item is the ability to recognize cause-and-effect relationships. This type of item usually contains two true propositions in one statement, and the examinee is to judge whether the relationship between them is true or false. The true-false item also can be used to measure some simple aspects of logic.

Advantage of True-False Item

A major advantage of true-false items is that they are efficient. Students can typically respond to roughly true-false items quickly in the time it takes to respond to multiple choice items. True-false items have utility for measuring a broad range of verbal knowledge and a wide sampling of course material can be obtained.

Limitations of True-False Item

Limitations of the true-false items are in the types of learning outcomes that can be measured. True-false items are not especially useful beyond the knowledge area. Another factor that limits the usefulness of true-false item is susceptibility to guessing. Successful guessing on the true-false item has effects that are at least as deleterious: First, the reliability of each item is low. Second, it has very little diagnostic value of such a test. Another concern that needs to be considered in the design of tests with true-false items is student response sets. A response set is a consistent tendency to follow a certain pattern in responding to test items.

Suggestions for Constructing True-False Items

The following are some common suggestions help true-false item writers to write true- false items: avoid trivial statements, avoid broad general statements, and avoid the use of negative statements, especially double negatives. When a negativeword must be used, it should be underlined or put in italics so that students do not overlook it. Avoid complex sentences. Avoid including two ideas in one statement, unless cause-effect relationships are being measured. Avoid using opinion that is not attributed to some sources, unless the ability to identify opinion is being specifically measured. Avoid using true statements and false statements that are unequal in length. Avoid using disproportionate numbers of true statements and false statements.

2.3.1.1.2. Multiple Choice Items

Haladyna, M. and Downing, M. (1989) states that the multiple choices item consists of two distinct parts: The first part that contains task or problem is called stem of the item. The stem of the item may be presented either as a question or as an incomplete statement. The form makes no difference as long as it presents a clear and a specific problem to the examinee. Second part presents a series of options or alternatives. Each option represents possible answer to the question. In a standard form one option is the correct or the best answer called the keyed response and the others are misleads or foils called distracters.

The number of options used differs from one test to the other. An item must have at least four answer choices to be classified as a multiple choice item. The typical pattern is to have four or five choices to reduce the probability of guessing the answer. A good item should have all the presented options look like probable answers at least to those examinees who do not know the answer.

Characteristics of Multiple Choice Items

John, P. (2004) suggests that multiple choice items are considered better than all items that can be scored objectively. The multiple choice questions is the most flexible of the objective type items. It can be used to appraise the achievement of any educational objectives that can be measured by a paper-and-pencil test except those relating to skill in written expression and originality. An ingenious and talented item writer can construct multiple choice questions to measure a variety of educational objectives from rote learning to more complex learning outcomes like comprehension, interpretation, application of knowledge and also those that require the skills of analysis or synthesis to arrive at the correct answer. Moreover, the chances of getting a correct

response by guessing are significantly reduced. However, good multiple choice items are difficult to construct. A thorough grasp of the subject matter and skillful application of certain rules is needed to construct good multiple choice items.

Advantages of Multiple Choice Items

Multiple choice item has the following advantages. It covers wide sampling of content, the problem or the task is well structured or clearly defined, has flexible difficulty level, efficient scoring of items, objective scoring, provide scores easily understood and transformed as needed: Multiple choice tests provide scores in matrices that are familiar to most of the score users i.e. percentiles, grade equivalent scores.

Limitations of Multiple Choice Items

The multiple choice items, despite having advantages over other items, have some serious limitations as well. It takes time to construct, it is susceptible to guessing and do not provide any diagnostic information.

Suggestion for Constructing Multiple - Choice Items

Abedi, J. (2010) suggests that be sure that the stem clearly formulates a problem. The stem should be worded so that the examinee clearly understands the question being asked before he reads the answer choices. Stem should be written either in direct question form or in an incomplete statement form. The stem of the item should present only one problem. Two concepts must not be combined together to form a single stem. Include as much of the item in the stem and keep options as short as possible: this leads to economy of space, economy of reading time and clear statement of the problem. Hence, include most of the information in the stem and avoid repeating it in the options. For example, if an item relates to the association of a term with its definition, it would be better to include definition in the stem and several terms as options rather than to present option in the stem and several definitions as alternatives.

Unnecessary words or phrases should not be included in the stem. Such words add to the length and complexity of the stem but do not enhance meaningfulness of the stem. The stem should be written in simple, concise and clear form. Avoid the use of negative words in the stem of the item. There are times when it is important for the examinee to detect errors or to know exceptions. When a negative word is used in a stem it should be highlighted. Use novel material

in formulating problems to measure understanding or ability to apply principles. Do not focus too closely on rote memory of the text that neglects measurement of the ability to use information. Use plausible distracters as alternatives. If an examinee who does not know the correct answer is not distracted by a given alternative, that alternative is not plausible and it will add nothing to the functioning of the item. Be sure that no unintentional clues.

The correct answer should appear at each position in almost equal numbers. While constructing multiple - choice item, some examiners have a tendency to place correct alternative at the first position. Some place it in the middle and others at the end. Such tendencies should be consciously controlled. Avoid using 'none of the above', 'all of the above', both a and b etc. as options for multiple choice questions. Alternatives should be grammatically consistent with the stem. Grammatical inconsistency provides irrelevant clue.

2.3.1.1.3. Matching Items

According to Gronlund, N. E. (1998) matching items consists of two parallel columns with each word, number, or symbol in one column being matched to a word, sentences, or phrase in the other column. Items in the column for which a match is sought are called premises, and the items in the column from which the selection is made are called responses.

Advantage of Matching Item

Gronlund, N. E. (1998) says that when you have a number of questions of the same type (homogeneous), it is advisable to frame a matching item in place of a number of similar multiple choice questions. Whenever learning outcomes emphasize the ability to identify the relationship between two things and a sufficient number of homogeneous premises and responses can be obtained, a matching exercise seems most appropriate. Hence it is suggested to use only homogeneous material in single matching exercise. e. g. inventions and inventors, authors and books, scientists and their contribution. The major advantage of the matching item is its compact form, which makes it possible to measure a large amount of related factual material in a relatively short time.

Limitations of Matching Item

Gronlund, N. E. (1998) states also though matching item has advantages, it has its own limitations. It is restricted to the measurement of factual information. Another limitation somewhat is the difficulty of finding homogeneous material that is significant from the viewpoint of our objectives and learning outcomes.

Suggestions for constructing matching items

Use only homogeneous material in a single matching exercise. Include an unequal number of responses and premises and instruct the student that responses may be used once, more than once, or not at all. Keep the list of items to be matched brief, and place the shorter responses on the right. Arrange the list of responses in logical order. Place words in alphabetical order and numbers in sequence. Indicate in the directions the basis for matching the responses and premises. Ambiguity and confusion will be avoided. And testing time will be saved. Place all of the items for one matching exercise on the same page.

2.3.2. Constructed Response / Supply Items

In the supply type items the question is so framed that the examinee has to supply or construct the answer on his own words. They generally include the following type: Essay, Short Answer and Completion.

2.3.2.1. Short-Answer and Completion Items

Heaton, J.B. (1988) says that the short-answer item and the completion item both are supply-type test items. Yet, they are included here for their simplicity. They can be answered by a word, phrase, number, or symbol. The short-answer item uses a direct question whereas the completion item consists of an incomplete statement. Short-answer item is especially useful for measuring problem-solving ability in science and mathematics. Complex interpretations can be made when the short-answer item is used to measure the ability to interpret diagrams, charts, graphs, and pictorial data. When short-answer items are used the question must be stated clearly and concisely. It should be free from irrelevant clues, and require an answer that is both brief and definite.

Advantages of Short-Answer Items

According to Heaton, J.B. (1988) short-answer test item is one of the easiest to construct. This is used almost exclusively to measure the recall of memorized information. In short-answer type item the student must supply the answer; this reduces the possibility that the examinee will obtain the correct answer by guessing.

Limitations of Short-Answer Items

It is not suitable for measuring complex learning outcomes. Unless the question is carefully phrased, many answers of varying degrees of correctness must be considered for total or partial credit. Hence it is difficult to score. These limitations are less troublesome when the answer is to be expressed in numbers or symbols, as in physical science or mathematics.

Suggestions for Constructing Short-Answer Items

According to Alderson, J.C. and et al. (1995) and Heaton, J.B. (1988) the following suggestions will help to avoid possible pitfalls and provide greater assurance that the items will function as intended. Word the item so that the required answer is both brief and specific. A direct question is generally more desirable than an incomplete statement. Do not take statements directly from textbooks to use as a basis for short-answer items. If the answer is to be expressed in numerical units, indicate the type of answer wanted. Blanks for answers should be equal in length and in a column to the right of the question. Do not include too many blanks.

2.3.3. Selecting Passages for Test Purpose

Heaton, J.B. (1988) suggests that care should be taken to select passages for which the selection is self-contained. Each passage should function as a whole, having a beginning, middle, and an end. Passages should be well-written and contain accurate facts. It is also best to avoid passages containing non Standard English. Passages should not come from textbooks or works that are frequently taught or highly anthologized. At each grade level, passages should represent a range of topics. Passages should be of probable interest to students at the target grade. Passage must contain enough information or ideas to serve as the basis for 10 to 12 test questions. Select passages from the internet only if they were previously published elsewhere as well. Passages must be of grade-appropriate length, depending upon the specifications of each program. The following general guidelines may be helpful.

Passages should be drawn from a variety of sources, including both magazines and books, unless the test specifications require otherwise. Do not locate passages exclusively from popular children's magazines. Passages should be selected to represent a variety of ethnic and cultural groups and include students from a variety of socioeconomic backgrounds. Passages should not contain content that suggests that the test developers are delving into personal values. Passages should not seek to advocate a particular value that may not be appropriate to all students taking the test.

It is crucial that passages be free from problems of bias or sensitivity. Content should show males and females, various ethnic and cultural groups, the elderly, and the disabled in a variety of positive roles. Passages should not contain words that might be demeaning to a particular group or references that might tend to stereotype.

2.3.3.1. Guidelines for Selecting of Specific Kinds of Passages

Narratives and Folktales Passages

As Heaton, J.B. (1988) says these type of passages must have a recognizable theme, round characters with sufficient motivation, literary value (imagery, developing characters, theme, and suitable style), suitable setting (safe and supervised for younger children), appropriate vocabulary, sufficient descriptive details, some conflict and action (conflict should be resolved at the end), an attention-getting beginning and a satisfying ending.

Personal Narrative Passages

These type of passage must have a recognizable theme, a character with sufficient motivation, literary value (imagery, developing characters, theme, and suitable style), suitable setting (safe and supervised for younger children), appropriate vocabulary, sufficient descriptive details and some conflict (either external or internal) and action (conflict should be resolved at the end) at attention-getting beginning and a satisfying ending.

Informational, Science, Biography, Persuasive, and Multicultural Passages

These type of passage must have an appropriate topic that is clearly presented, connected thoughts with specific support, a strong sense of order (chronological, cause/effect, comparison/contrast, classification or division,) where appropriate, appropriate vocabulary, focus, and accurate details for commissioned pieces: two supporting documents (no more than one from the

world wide web) with information used underlined and coded to textinteresting and compelling information, information focused on specific concepts, not scattershot
“How to” articles must have detailed and specific informationappropriate sequential ordera list of materials, where appropriate vocabulary and concept for the grade levelexcellent organization.

2.4. Principles for Designing Effective Classroom Tests

Teachers who design their own tests can convert ordinary, traditional tests and test situations into authentic, intrinsically motivating learning opportunities designed for learner’s best performance and optimal feedback. According to Lourdes (2008/09), the following principles help language teachers to design effective classroom tests for their students:

Principle 1: Strategies for Test-takers

Lourdes (2008/09) suggests that teachers need to offer their learners appropriate& useful strategies for taking the test. Because through strategies-based test-taking, learners can avoid miscues due to the format of the test alone and they should also be able to demonstrate their competence throughan optimal level of performance. The following points should be considered by language teachers before, during, and after the test:

Before the test:

1. Give students all the information you can about the test. Exactly, what will the test cover?Which topics will be the most important? What kind of items will be included? How long will it be?
2. Encourage students to do a systematic review of material.
. E.g. skim the textbook & other exam material, outline major points, write down examples, etc.
3. Give them practice tests or exercises, if available.
4. Facilitate formation of a study group, if possible.
5. Caution students to get a good night’s rest before the exam.
6. Remind students to get to the classroom early.

During the test:

1. As soon as the test is distributed, tell students to quickly look over the whole of it so as to get a good grasp of its different parts.
2. Remind them to mentally figure out how much time they will need for each part.

3. Advise them to concentrate as carefully as possible.
4. Alert students a few minutes before the time is over so that they can proofread their answers, catch careless errors, and still finish on time.

After the test:

1. When you return the test, include feedback on specific things the student did well, what they did not do well and if possible, and the reasons for such judgments on your part.
2. Advise the students to pay attention when you discuss test results in class.
3. Encourage questions from the students about test format and test results, scoring procedures, etc.
4. Advise students to make a plan to pay special attention in the future to points that they were weak on.

Sometimes, students don't know exactly what is being tested when they take a test. Sometimes they feel fear, for a variety of reasons, that the test isn't testing what it is supposed to test.

Principle 2: Face validity

Lourdes (2008/09) says that sometimes, students do not know exactly what is being tested when they take a test. Sometimes they feel anxiety for a variety of reasons, that the test is not testing what it is supposed to test. Face validity means that in the students' perception, the test is valid. Teachers can help foster this perception with a carefully constructed, well-thought-out format; a test that is clearly doable within the allotted time limit; items that are clear and uncomplicated, directions that are very clear; tasks that are familiar and relate to their text book; a difficulty level that is appropriate for students.

Principle 3: Authenticity

According to Lourdes (2008/09) teacher need to make sure that the language in their test is as natural & authentic as possible (e.g. instead of making up examples, look for them on the Internet and, if necessary, adapting them to needs of the teachers and student level, etc.) Try to give language some context so that items aren't just a string of unrelated language samples. The tasks themselves need to be tasks that the students have practiced and feel comfortable with. (That is a test is not the time to introduce brand-new tasks, for the teacher won't know if student poor performance is a factor of the task itself or of the language the teacher are testing).

Principle 4: Beneficial Washback

According to Lourdes (2008/09) beneficial wash back is the benefit that tests offer to students' learning. When students take a test, they should be able, within a reasonably short period of time, to utilize the information about their competence that test feedback offers.

In other words, formal tests must be learning devices through which students can receive a diagnosis of areas of strength & weakness; their incorrect answers can become windows of insight about further work; your prompt return of written tests with feedback is fundamental to intrinsic motivation.

Teachers can enhance beneficial wash back by providing a generous number of specific comments on test performance. Teachers need not always to think their job is done by just returning tests with a letter grade or number score on them: they give absolutely no information of intrinsic interest to students. At best, they give a relative indication of a judgment of performance as compared to others in the class –which fosters competitive, not cooperative, learning. Give praise for strengths together with constructive criticism of weaknesses. Give strategic hints on how a student can improve certain elements of performance.

2.5. Practical Steps for Test Construction

Step 1: Test toward Clear Objectives

According to Madsen, H. (1983) a test objective must be as clear as crystal to you what it is that you want to test. Don't just give a test because it is Friday, or 3rd week of the course, or for students to have something to do in class. Instead, make a list of everything you think your students should know or be able to do base on the material the students are accountable for. Your objectives can be, for testing purposes. Such as grammar, tag questions, Simple past in negative sentences & information questions, Irregular past tense verbs, Conjunctions, Communication skills, Guessing what happened, Finding out who did something, Giving reasons, Asking for confirmation.

Step 2: From Objectives, Draw up Test Specifications

According to Madsen, H. (1983) says that ‘don’t freak!’ Test specifications for classroom use can be a simple & practical outline of your test. For the above objectives (step 1) your specifications will indicate, how you will divide up the 40-minute test period or what skills you will test and what the items will look like.

These informal classroom-oriented “specs” give you an indication of: which of the objectives you will cover, what the item types will be, how many items will be in each section, and how much time is allocated for each. Beware! You may decide not to test everything you teach or not to include all the language skills in your test for reasons of practicality.

Step 3: Draft Your Test

The first draft will give the teacher a good idea of what the test will look like, how students will perceive it (face validity) the extent to which authentic language and contexts are present.

Step 4: Revise Your test

Madsen, H. (1983) suggests that at this stage, the teacher will work through all the items he has created and ask himself a number of important questions: Are directions to each section clear? Is there an example item for each section? Does each item measure a specified objective? Is each item specified in clear, simple language? Does each multiple choice item have appropriate distracters, i.e. are the wrong items clearly wrong & yet sufficiently alluring that they are not ridiculously easy? Does the difficulty for each item seem appropriate for your students? Do the sum of the items & the test as a whole adequately reflect the learning objectives?

Step 5: Final-Edit and Type the Test

Madsen, H. (1983) suggests that ideally, the teacher would try out all his tests before actually administering them. In real classroom practice, the try-out phase may be virtually impossible. In his final editing of the test, before typing it for presentation to his class, imagine he is one of his students and goes through each set of directions and through all items slowly & deliberately, timing himself so as not to underestimate the time students will need to complete the test. If he needs to shorten or lengthen the test, he makes sure the test is neat & uncluttered on the page, reflecting all the care & precision he has put into its construction.

Step 6: Utilize your feedback after administering the test

After you give the test, you will have some information about how easy or difficult it was; the time limits; student affective reaction to it, & their general performance. Take note of all this feedback & use it for making your next test better than the previous ones.

2.6. Conclusion

The review literature of this chapter has attempted to describe literatures relevant to teachers' practice of designing English language tests at classroom level. Accordingly, it has tried to give an overview of what has been said by different scholars around this research topic. Therefore, the literature of this chapter has been organized in discussing effective test item construction techniques, items commonly used for classroom language tests together with their advantage, limitations and suggestion for constructing them. Effective principles for English language teachers in designing classroom tests and practical steps for classroom test construction have been discussed under the chapter in line with the objectives of the study.

Chapter Three

Research Methodology

This study intended to examine teachers' practice of designing and implementing English language tests at classroom level. In order to study teachers' practice of designing and implementing English language exam at classroom level, the research methodology; thought to be appropriate to this study was designed in line with the objectives and research questions stated under chapter one of this thesis. In this chapter, the research design and procedures employed to carry out this study were presented. First, the research design employed in conducting the study was described. Next the research site, population of the study and sampling were stated, followed by the instruments used to collect data. After that, procedures of data collection and procedures of data analysis were presented consecutively.

3.1 Research Method

Mixed research method was used in this study. The method was selected because researchers in the field comment that complex educational phenomenon that needs to be examined through employing different research approaches should employ the mixed research method. In light of this, Dornyei (2007) reports that in applied linguistics, we find many studies that have combined methodologies. For instance, as Magnan (2006) cited in Dornyei (2007) reports that over the 1995-2005 period 6.8 percent of the research papers appearing in the modern language journal used mixed research methods, which is relatively high when it is compared to the total number of qualitative studies.

As Lazaraton (2005:219) cited in Dornyei (2007) concluded that researchers in the field of applied linguistics would better see more studies that combine qualitative and quantitative research methods, since the strength of one method can be utilized to overcome the weakness of another method used in the study. Furthermore, according to Dornyei (2007) mixed research method has better advantages over single type of research method. Because, mixed research method helps to analyze multi-level complex issues. For instance, it helps us to get a better understanding of a complex phenomenon by converging numeric trends from quantitative data from qualitative data. Words can be used to add meaning to numbers and numbers can be used to add precision to words. It has also a unique potential to produce evidence for the validity of research outcomes through the convergence and corroboration of the findings.

Cheng et al., (2004) reported that studies in the area of complex educational phenomenon need mixed methods research to get reliable and valid data. In addition to this, Bailey (1999) recommends that educational studies should employ methods that investigate what teachers and students perceive and what they practice. This shows that combining both quantitative and qualitative methods can be useful for this study. Therefore, to take the advantages of the strengths of the two methods, the mixed research method was used in this study. However, the research used mostly the qualitative method rather than the quantitative method as it was inclined more toward qualitative data.

3.2. Research Site

The study was conducted in Shirka High School, Arsi Zone, Oromia Regional state, grade 10 English teachers in focus. The site was selected based on the following scholars' suggestions.

According to Blaxter et al (2006) the cost of the research should not be forgotten by the researcher. Unless he has an employer, funder or sponsor who is going to meet absolutely all of the costs of his research project he should be aware of the different costs associated with alternative kinds of his research. For even a relatively modest project, the costs which he may have to bear must be a small amount.

Lee, R. (2000) states that virtually any research project involves questions of access to resources like people, places and/or documents. Access can be seen as relating as much to the resources researcher has available (e.g. a good library), as to the subjects of his research. Obviously, from this point of view, it may make sense to choose a topic for which he believes access will be less problematic. This may suggest doing his research project within his institution, though that does not guarantee there will be no problems.

Berry, R. (2004) suggests that in choosing a research topic, it makes sense to think about the methods you will use to collect and analyze data as soon as possible. If you have a choice, consider the methods you will enjoy using or not. For example, if you like talking to people, you might be well advised to make use of interview methods. If, on the other hand, you don't like talking to people, you might think about undertaking library or document-based research. The methods you use are a key part of your research, so you need to understand something of the alternatives available to you, and their strengths and weaknesses. If

you enjoy or have a flair for a particular method, this can make your research project more interesting, and help to motivate you to carry it through. Or you might like to use your research project to learn about, or develop your skills in, methods you are not familiar with.

Therefore, based on the above scholars' suggestions by considering the cost of research, accessibility to resources, methods used to collect data and analyzing data this site was selected for the study.

3.3. Population and Sampling

The sample is the group of participants whom the researcher is actually examines in an empirical investigation and the population is the group of people who the study is about (Dornyei, 2007). Accordingly: the target populations of the study were all English language teachers in Shirka high school. There were totally ten English language teachers in the school. Out of the total number of ten teachers, one was female and the rest nine were males. Because their number was convenient and manageable to gather data all of those English language teachers were included in the sample size of the study using comprehensive sampling technique. The primary purpose of selecting teachers as participants was to get reliable data through semi-structured interview about their practice of designing tests since the problem studied was directly related with their work.

3.4. Instruments of Data Collection

According to Holsti (1968, P-608) definition content analysis is "any technique for making inferences by systematically and objectively identifying special characteristics of messages of written documents or recorded verbal communications. Selltize et al (1959) suggest that quantitative content analysis result in somewhat arbitrary limitation in the field by excluding all accounts of communications that are not in the form of numbers as well as those that may lose meaning if reduced to a numeric form (definitions, symbols, detailed explanations, and so forth). Other proponents of content analysis, notably Smith (1975) suggests some blend of both qualitative and quantitative analysis should be used. Smith (1975, P.218) explains that he has taken this position "because qualitative analysis deals with the forms and antecedent-consequent patterns of form, while quantitative analysis deals with duration and frequency of form." Since document analysis and semi-structured interview are written document and

recorded verbal communications respectively, the researcher used mixed- methods research which includes both qualitative and quantitative methods in his study based on the above suggestions.

Therefore, based on the above selected research method, the researcher used two instruments to collect data. The main data collecting instrument used by the researcher was document analysis (teacher-made English language final examinations). This instrument was selected to see the actual teachers' practice in designing and implementing their exams at classroom overtly. To analyze the documents two types of checklists namely checklist for multiple choice and checklist for teachers' techniques in selecting reading passages were set from different sources such as strategies, techniques, principles and suggestion written regarding test designing activities. During checklist preparation, special care was taken to avoid the mismatch between the objectives of the study and items of the checklists.

3.4.1. Document Analysis

As it is stated above, the main instrument used in the study was document analysis. In this case, first semester teacher-made classroom exams were analyzed to see whether or not the teachers design the test items following strategies, principles and essential guidelines for classroom test preparation. In analyzing the exam papers, checklists (see Appendix A) were set based on the principles, techniques and strategies teachers should use in designing classroom tests. Teacher made first semester final exams of the academic year 2013-2015 were collected from Shirka high school teachers. The collected exams included only first semesters of 2014—2015 years. In another word, the three consecutive semesters were collected since teachers prepare only first semester exams. The teachers prepare only first semesters' exams because of that the second semester final exams are prepared at national level as model exams which the teachers use as second semesters' final exams in their school. The sole purpose to collect the exams was to analyze the actual teachers' practice in their own tests. So the researcher used data obtained from collected exams to cross-check with the interview responses obtained from teachers to study the problem.

3.4.2. Teachers' Interview

The other instrument used in this study was teachers' semi-structured interview. Interview is important data gathering instrument in applied linguistic qualitative study. The essentiality of interview for such applied linguistic study has been stated by different scholars. For instance, as (Block; 2000,Rubio;1997;Richards 2003) cited in Dornyei(2007), interview is a known instrument that works so well as better instrument although there is a range of qualitative research techniques available for researchers. Based on the scholars' comment, semi-structured interview was designed and conducted with ten teachers to check their awareness about teacher-made classroom exams, techniques they use to design tests at classroom level, principles they follow in designing and implementing tests at classroom levels and test items they often use. The interview questions were written by searching information from different literatures and course offered in universities regarding testing.

3.5. Procedures of data collection

The data collection process was carried out as follows. First, three consecutive semester's classroom final exams set by English language teachers at the school level were collected and read critically by the researcher in order to realize the strong and weak side of teachers practice in designing tests. This was done to check whether the preliminary problems assumed by the researcher existed or not. After the exams were read deeply, checklists were prepared for the exams analysis and interview items were set by searching information from different sources of related literatures and course's notes. Two types of checklist items were set. The first type of checklist items was prepared to analyze multiple choice items and the second type was prepared to analyze teachers' techniques in selecting reading passages for their exams. Following checklist preparation semi-structured interview questions were also prepared for English language teachers. After checklists and interview items preparation were completed, data were collected from teachers-made exams using both types of checklists mentioned above. Moreover, to strengthen and make the finding more reliable and to get new generated idea, semi- structured interview was also held together with ten teachers. The interview was conducted with ten teachers for two days: from March, 25-26/03/2015. It took about twenty-five to thirty minutes for an interviewee to respond to the questions.

3.6. Procedures of Data Analysis

First, to analyze data intended to be obtained from document analysis or teachers' made exams, each semester's exam was read seriously in line with items developed in checklists. When the problems appeared in the exams paper, immediately signs were marked in the table where problems were appeared in both multiple choice items analysis table and check list table for selecting passages. After the problems were identified in the tables, specific interpretation was given for each of the items in the table where problems appeared. Finally, the specific information found from both tables were brought together and generalization regarding checklist analysis was written. Furthermore, the interview responses given by interviewees were analyzed at individual levels to compare and contrast the responses given by each teacher through nominating the interviewees as teacher-1, teacher-2, etc... for all ten respondents (see appendix C). After individual teachers' responses analysis had been completed, the final interpretation was given regarding interview response analysis.. This was done, to minimize the influence of one component on the role of the other and to obtain substantiated findings. In this regard, Dornyei (2007: 245) recommends the following:

... The analysis of the data should proceed independently for the qualitative phases and mixing should occur only at the final interpretation stage. in many cases, it may be better to keep the analyses separate and only mix the qualitative results at a late[r] stage to illuminate and corroborate each other.

Finally, the data obtained from interview and document analysis (exam paper analysis) were analyzed at final stage. Supporting this, Creswell, (2010), notes that the mixing of the qualitative data analyses findings is usually done at the interpretation or discussion section. Thus, the analyses of this study's data were made in a separate and parallel manner as recommended by these scholars.

Chapter Four

Data Analysis and Discussion

This chapter presents the data gathered through checklists for analyzing teacher-made tests and semi-structured interview. These tools were used because it was hoped that the result obtained using them would be valid. The discussion of the data collected using the above instruments were organized into two major sections. The first section presents results obtained from test items analysis and the second section presents results obtained from semi-structured interview made with English language teachers. Here, the results obtained from test items analysis and structured interview made with English language teachers have been analyzed and discussed. In analyzing the results obtained from test item analysis, checklists for multiple choice items and passages selection were used consecutively. The checklists' items were taken from suggestions, principles and strategies suggested by different scholars in the English language testing literature and courses offered on language testing in the universities. On other hand, the semi-structured interview items were constructed by focusing on teachers' awareness about strategies, techniques, and principles of designing teacher-made tests. Therefore, based on the data obtained through the two instruments, the following analyses and discussions were made.

4.1. Analysis and Discussion of Test Items

The following two tables contained items constructed from suggestions, principles, strategies and techniques suggested by scholars to design effective language exam at class room level as well as from courses offered regarding testing. The items were designed in helping to check how teachers write exam items. The primary focus of the checklist points were to point out the limitations teachers had in designing exams at classroom level. Thus, the checklist points were used to analyze consecutive three years' teacher-made exams (exam prepared at department level). Finally, the exams were cross-checked with checklist items in the table and the results were obtained.

4.1.1. Analysis and Discussion of multiple choice items

The data that were collected from documents or first semester teacher-made English exams were presented in the table. In the table checklist points were written based on the three aspects of multiple choice writing criteria. These were general construction criteria of the questions, construction of the stem portion of the questions and construction of the options of the questions.

Table 1: Multiple Choice Items analysis

No	Checklist Points	Exam Year						Remark
		2013		2014		2015		
		Yes	No	Yes	No	Yes	No	
	a. General Construction criteria of the questions							
1.	The teacher distributed correct responses among the alternative positions.		✓	✓		✓		
2.	As, Bs, Cs, and Ds have nearly equal proportion as the correct answer.		✓	✓		✓		
3.	The teacher kept the number of alternatives at five and four.	✓		✓		✓		
4.	The teacher used negatively stated items in limited amount.	✓		✓		✓		
5.	The negatively stated items used were underlined and highlighted well.		✓				✓	
6.	There was not more than one best answer to the stem.		✓	✓		✓		
	b. Construction of the							
		Exam Year						Remark
		2013		2014		2015		

No	Checklist Points	Yes	No	Yes	No	Yes	No	
7.	The teacher used direct question rather than included as part of the stem, so that all options can be clearly distinguished.		✓	✓			✓	
8.	The teacher included much of the item in the stem and kept the alternatives short and brief.	✓			✓		✓	
.	<i>C. Construction of the options of the questions</i>							
10	The teacher kept all the alternatives in a similar format. (i.e. all phrases, all words, all sentences.)		✓		✓		✓	
11	The teacher listed options on separate lines rather than included as part of the stem so that all options can be clearly distinguished.	✓		✓			✓	
12	All options were plausible responses to the stem?		✓	✓		✓		
13	The teacher didn't include poor alternatives for the sake of having more options.		✓	✓			✓	
No	Checklist Points	Exam Year						Remark
		2013		2014		2015		
		Yes	No	Yes	No	Yes	No	
14.	All of the options had grammatical consistency with the stem (tense, number,		✓	✓			✓	

	etc...)							
15.	All the alternatives were written approximately in the same length.	✓	✓	✓			✓	
16.	Capital letters (A. B. C. D.E) were used as response signs rather than lower case letters.	✓		✓		✓		

The above table shows data obtained from checklists made up of general criteria for item construction, stem portion, and option of multiple choice item type. Based on these checklists grade ten English language teachers' practice of designing multiple choice items for their students were analyzed. Accordingly, the discussion focused on problems exhibited in the consecutive three years first semester final exams prepared at department level by these teachers. In line with the criteria of the checklists, the prominent problems in the exam were discussed as follow. With regard to general construction of the items, few problems were seen in 2013 exam regarding the way the teachers distributed correct answers among the alternative positions (See Appendix C, P.4).

As it can be seen similarly in item five (5) of this section, in the three years, some of the negatively stated items used were not underlined and highlighted. In addition to this, item six of this section revealed that there was one best answer to the stem in the year 2013. Regarding construction of the stem portion of the question almost many of teacher's problems in designing multiple choice items seem to be revealed well. As item seven (7) showed in the consecutive three years exams the teachers seemed not to use the stem portion of the question to present the problem or question as clearly as possible (See Appendix C,Q.5,6,10 in 2013, Q.11 in 2014 and Q.13 in 2015). In item eight (8), it was seen that the teachers did not include all parts of the stem in a direct questions and all options could not be distinguished clearly.

In the same way, in item nine (9) it could be observed that the teachers did not include much of item in the stem and kept the alternatives short and brief in years 2014 and 2015. The problems could be seen in some questions on different pages (See Appendix C:Q.13, 17 P.2-3 in 2014 and Q.10, 18, 19 P.1-2, in 2015).

As the information displayed in item ten (10) shows, the teacher did not keep all the alternatives in a similar format. These problems were seen in the three years. The most prominent problem could be seen especially in the last two years. (See Appendix C: Q.9 P.2 in 2013, Q. 3, 12, 13 in 2014 and Q.3, 16, 32, 34, P.1-4 in 2015).

As far as item eleven (11) concerned with teachers' practice of listing options on separate lines rather than including as part of the stem to make all options to be distinguished clearly, few problem could be seen in the year 2015. In the same manner in item thirteen regarding including poor alternatives for the sake of having more options, no more problems could be seen except few of them in the year 2013 and 2015.(See Appendix C: Q.22 P.3 in 2015).

Regarding item fourteen (14) that dealt with the grammatical consistency of options with the stem, some problem were appeared to be seen in different questions on different pages in the year 2013 and 2015. In the same, in item fifteen (15) concerning the length of the alternatives few problems could be observed. (See Appendix C: Q.3, 7, 8, 2, P.1-3 in 2015).

In general, the above checklist of multiple choices item contained three sections entitled as general construction criteria of the questions, construction of the stem portion of the question and construction of the option. The analysis of the data revealed some problems English language teachers have in constructing multiple choice items and the gap between some suggestions of multiple choice items construction provided by scholars. As a result the most problems seen were; negatively stated items were not underlined and highlighted, some stems were not as clearly as possible to present the question.

4.1.2. Analysis and discussion of Teachers' Techniques in Selecting Reading Passages.

The data that were collected from documents or first semester teacher-made English exams were presented in the table. In the table checklist points were written based on the criteria that an English language teacher should follow when he writes passage for exam purpose. Based on the data gathered using these checklists, the analysis and discussion of the data were presented below the table as follows.

Table 2: Analysis of Teachers' Techniques in Selecting Reading Passages

No	Checklist Points	Exam Years						Remark
		2013		2014		2015		
		Yes	No	Yes	No	Yes	No	
1.	The passage has a beginning, middle and end.	✓		✓		✓		
2.	The passage is written well in standard English containing facts.		✓	✓		✓		
3.	The passage was come from other sources apart from text book.	✓		✓		✓		
4.	The passage could be interesting and to the level of students target grade level.		✓	✓		✓		
5.	The passage’s length is the appropriate to the grade level (300-1200 words).	✓		✓			✓	
6.	The teacher presented the passage that includes students from variety of cultural back grounds.	✓			✓		✓	
7.	The passage’s content was free from unnecessary thoughts that burden works	✓			✓		✓	

No	Checklist Points	Exam Years						Remark
		2013		2014		2015		
		Yes	No	Yes	No	Yes	No	
8.	The teacher wrote the passage that advocate common value to all students taking the test.	✓			✓		✓	
9	The passage is free from problems of bias or sensitivity (i.e. avoid words that demean a particular group or references that might tend to stereotype.)	✓		✓		✓		
10	The teacher sited the source of the passage. (Author’s name, year of publication, place of publication...etc)		✓		✓		✓	
11	The passage is free from over using idiomatic expressions or figurative language unnecessarily to students whose first language is not English.	✓		✓		✓		
12	The teacher avoided topic in his passage that is likely to upset, or unduly distract students and affect their performance on the test.	✓		✓		✓		

The above table shows data obtained from analysis of teachers' techniques in selecting reading passages. Each item was taken from the literature related with techniques that English teachers should use when they select reading passages for exam purposes. Therefore, based on data obtained using checklists the following discussions were made carefully.

Item 2 showed that in 2013 the passage was not written well in standardized English. First of all the passage was not written using normal font size. In addition, the passage has some difficult

phrasal verbs like "pleaded with" was not to the level of the students. According to education testing service (2002) passages for grade 9-10 must be 300 to 1200 words, but this was not seen in the passage. Because it contained words not more than one hundred fifty. (see Appendix D). Thus, the passage seems to have limitation to test students' deep comprehension ability. Furthermore, item ten (10) showed that in 2013, 2014 and 2015, the teachers did not cite the source of the passage where they had taken the passages. In the same way, item eleven (11) in 2013 showed that the teachers selected passage with some difficult phrases such as 'pleaded with' ---ground nuts to carry---

To sum up, the analysis of data obtained from checklists of teachers' techniques in selecting passage appeared to show some problems English language teachers had when they selected passage for exam purpose. The most problems seen were inability to give focus on language difficulty level, not to cite source of the passage, the passages were not written legibly and the other contained difficult concepts.

4.2. Analysis and Discussion of Teachers' interview Responses on each question.

The analysis of data obtained from semi-structured interview conducted with ten interviewees was presented as follow. The analysis of the data was categorized into four sub-titles based on the four basic research questions of the study. These were analysis and discussion of interviewees' responses regarding strategies and techniques teachers use in designing classroom tests, principles that English language teachers follow in designing English language tests at classroom level, type of test items frequently used by English language teachers and factors that affect teachers' practice in designing effective tests at classroom levels.

4.2.1. Analysis and discussion of teachers' responses regarding strategies and techniques they use in designing classroom tests.

As it can be understood from the interviewees' responses, all of the interviewees responded that they give final exams to grade for ten students twice within a year. Of the two times exams, they take part only in designing the first semester exams one. As it is understood from the reason they gave, they prepare only the first semester exam. They do not prepare the second semester final exam. They use the model exam that is prepared at national level instead of second semester final exam (see Appendix D, Q.1 responses).

As it can be analyzed from the responses given by all the interviewees, most of the interviewees said that they have reasons to give exam to their students. Almost their reasons to give exams for their students revolve around evaluating what students had already learnt. But only few interviewees responded that they have reasons to give exams depend on their own need. Regarding this point, teacher2 responded that he gives final exam based on his need. He said that gives students exam when he wants to assess students' total achievement about the subject matter they had already learnt.(See Appendix D,Q.2)

According to the responses given to the third question, most of the interviewees believe in about the importance of giving exam for students. All the interviewees responded that their exams have great importance for their students for different purposes. From the responses they gave. It can be deduced that they give their students exam to evaluate students academic achievements and to enable students to be benefited from the exam: by evaluating their weakness and strength on the content area, by looking back to the content area where they have difficulties, and by being motivated to study more. But few of the interviewees responded that the exam helps students to show the way national exam is prepared and to encourage them study more to promote from grade to grade (See Appendix D,Q.3). Speaking generally, it seems that most teachers have concept about the merits of their exam to their students.

Mainly the responses presented by the interviewees expressed that what the teachers think as techniques to design exams at class room level are some activities they do at department level when they design the exams, but which are not the actual techniques that must be implemented in designing the exams (See Appendix, D.4). Similarly all of the respondents declared that few item types are shared for teachers. Then, a teacher constructs types of items he or she got. At end they come to the department and decide on the exam items. On other hand, few of them could tell some effective techniques suggested by scholars, such as planning the exam, writing exam objectives, out lining the test contents, decision and advice of subject matter experts or colleagues, preparing the test according to the plan. Test administration and use, scoring the test and reviewing/analyzing test results. This shows that the teachers did not use the right techniques help them to design class room exams.

From the interviewee's responses it can be realized that writing exams plan before constructing exam items are not exercised. All the ten interviewees responded that they do not use any written plan personally and at department level. Planning the test which includes several activities that language teachers need to carry out before constructing new exam such as drawing up test 'blue print'. Specifying the content and objectives of the test, type of items and time limit are seemed to be ignored by the teachers. (See Appendix D, Q.5).

4.2.2. Analysis and discussion of teachers' responses regarding principles that they follow in designing English language tests at classroom level.

As it can be observed from the respondents answers, most of the interviewee appeared to believe in the importance using activities such as writing objectives of the exam, outlining exam content have their own use for teachers. However, no one of them could tell the practical implementation of these techniques in their school. (See Appendix D, Q.6)

It is possible to generalize that most respondents listed activities that they do before distributing the exam papers for their students. As their responses almost revealed, they arrange seating, avoid noise sound, checking questions on all pages, tell students to fill necessary information, tell to look at each instruction, and advise to concentrate as carefully as possible on question and alert students mentally how much time they use for each item type (See Appendix D, Q.7). On other hand, important activities to be done after exams were scored such as: comparing students' test results by assigning ranks to the score, interpreting the individual result performance with the whole group that has taken the test, returning and reviewing test results with the students, giving feedback on the performance to motivate students for improvement and encourage to learn from their mistakes, did not mentioned largely.

Noticeably majority of the interviewees were inclined to answer some factors that can affect students' psychological emotion during they take their exam in the classroom. As their responses revealed, students emotions can be affected in the classroom by invigilators behavior, speed their peers use on doing exam, fear of exam, vague exam instructions and students' personal problems (See Appendix D, Q.8). Out of all respondents those who were interviewed majority of them appeared to give relatively the same responses. They seemed they are not sure enough that their

students exactly know the content areas their exams cover and only few of them seemed to get help their students know the content areas of the exams. (See Appendix D, Q.9).

Except few of the interviewees, most of them were inclined toward taking care of the language difficulty when they prepare exams for their students. Somewhat it seems that they usually use simple languages to construct exam items (See Appendix D, Q.10).Relatively, a lot of responses that were given by the interviewees would seem to have less close relations with some of the suggestions provided in the review of literature part of this study paper. However, few of the responses do not have relation with the questions. In general, most of the teachers have no clear reasons to revise their exam items with their students(See Appendix D, Q.11).

4.2.3. Analysis and discussion of teachers' responses regarding type of test items they frequently use in designing class room tests.

As it can be estimated from the responses given by the interviewees, most of the teachers seemed to use multiple choices item types repeatedly than the other types of test items. As the majority of reasons provided from the interviewees tends to show, the primarily reasons for teachers to use multiple choice usually is to relate teacher-made item types with those of national exams and to get students practice in scoring better results in national exams. (See Appendix D, Q.12).

4.2.4. Analysis and discussion of teachers' responses regarding factors that affect their practice in designing effective tests at classroom levels.

Out of all interviewees those interviewed on the question of their practice in running training on test designing techniques, a majority of them reported that they have not taken training on test item designing practice at school level in their department and have no experiences of saving exams in exam bank and storing in special area. This seems that there is limitation with these teachers in exchanging good practice of designing language tests among themselves.(See Appendix D, Q.13).

Out of all interviewees, those, with whom interview was made, almost all of them appeared to responded that they have never taken any training regarding test designing practice at district education offices, regional education offices and ministry of education level to renew their experiences of test designing skills except the one some of the took as measurement and evaluation course at higher institutions (See Appendix D, Q.14)

4.3. Discussion of Teachers' overall Interview Responses

As it can be understood from the interviewees' responses regarding their awareness about exam delivering time, all of the interviewees responded that they give final exams to grade ten students twice a year. They take part only in designing the first semester one, but they do not prepare the second semester final exam. They use the model exam that is prepared at the national level instead of second semester final exam. On the other hand, regarding the teachers' awareness about their exam purposes, most of the interviewees responded that they have reasons to give exam to their students. Almost their reasons to give exams for their students revolve around evaluating what students have already learnt.

According to the responses given by the interviewees regarding the importance of their exams to their students, most of the interviewees believe in about the importance of giving exam for students. Most of the interviewees responded that their exams have great importance for their students for different purposes such as to evaluate students' academic achievements and to motivate students toward learning. As a result students can be benefited from the exam by evaluating their weakness and strength on the content area and go back to the content area to read deeply. Some of the interviewees responded that the exam helps students to show the way national exam is prepared and to encourage them study more to promote from grade to grade. Generally, it seems that most teachers have awareness about the merits of their exam to their students.

Regarding teachers' experiences in using test techniques that help them to design exam items. They reported that they did not have special test designing techniques. What they reported as common technique they use to design test items is sharing types of test items and writing them individually at home and bringing to the department to discuss on the items correctness. In connection to their experiences of using some effective techniques suggested by scholars, they do not use techniques such as planning the exam, writing exam objectives, outlining the test contents, decision and advice of subject matter experts or colleagues, preparing the test according to the plan, test administration and use, scoring the test and reviewing or analyzing test results.

Relatively responses given regarding teachers' practice of using written plan, it could be realized that writing plan before constructing exam items are not exercised by teachers. Almost the majority of interviewees responded that they do not use any written plan individually and as department members. Planning the test which includes several activities such as draw up test 'blue print', specifying the content and objectives of the test, type of items and time limit that language teachers need to carry out before construct the new exam seemed to be ignored by the teachers.

As it can be observed from the interviewee's responses, most of them reported that they believe in the importance of using activities such as writing objectives of the exam, outlining exam content have their own use for teachers. However, no one of them responded that they knew the use of the above techniques theoretically but they did not use them practically. Most of the interviewees listed activities that they do before distribute the exam papers for their students. As their responses revealed, they arrange sitting, avoid noise sound, checking questions on all pages, tell students to fill necessary information, tell to look at each instruction, and advise to concentrate as carefully as possible on question and alert students mentally how much time they use for each item type. On other hand, important activities to be done after exams were not mentioned largely by the interviewees.

Noticeably majority of the interviewees were inclined to answer some factors that can affect students' psychological emotion when take exam in the classroom. As their responses revealed, students emotions tend to be affected in the classroom by invigilators behavior, speed their peers use on doing exam, fear of exam, vague exam instructions and students personal problems. Regarding the question how the teachers help their students to know the content area of the exam, majority of them appeared to give relatively the same responses. They seemed not to be sure enough that their students exactly know the content area exams cover and only few of them reported that they help their students know the content areas of the exams by telling the area orally.

With regard to types of items repeatedly used by the teachers in their exam, most of them responded that they use multiple choices than the other types of test items. As the majority of interviewees tended to explain the primarily reasons for teachers to use multiple choice

repeatedly in their exam is to relate teacher-made item types with those of national exams and to get students practice in scoring better results in national exams as primarily purpose. On other hand, some of the teachers reported that they use multiple choice item types because it is easy to construct, easy to score and save time to construct it.

Concerning teachers' practice of taking and giving training on test designing, majority of them reported that they have not been given training on test item designing practice at department level. They have no experiences of saving exams in exam bank and storing in special area. This seems that there is limitation with these teachers in exchanging good practice of designing language tests among themselves through in-staff training and store exams to use in the future.

Finally, regarding teachers' participation of in-service training on test designing, almost all of the interviewees responded that they have never taken any training regarding test designing practice at district education offices, regional education offices and ministry of education levels to update their experiences of test designing skills except some training they took in measurement and evaluation course when they were in higher institutions.

Chapter Five

Summary, Conclusions and Recommendations

5.1. Summary

The aim of this research paper was to assess English language teachers' practice of designing classroom tests in grade ten. To study the problem systematically, four specific research questions set out of one general objective were formulated to address the issue. In line with the specific objectives designed, four research questions were prepared. Furthermore, to enable to deal with the questions; related literatures to the addressed topic and two data gathering tools namely; document analysis and semi-structured interview were selected. Based on selected research tools, data were collected from two sources; from document or teacher-made exams using checklists and from English language teachers through semi-structured interview. The collected data were first analyzed and interpreted at two levels; at checklist analysis and interviewees' responses analysis. Based on this the major findings of the study were summarized as follows:

1. Regarding techniques teachers use to design their exams at classroom level:

- Most interviewees reported that they have no common techniques that they use in designing exam items at department level. Similarly all of the respondents explained that few item types are shared for teachers. Then, teachers construct types of items they get. At end they come to the department and decide on the exam items. On other hand, most of them do not use some effective techniques suggested by scholars such as planning the exam, writing exam objectives, outlining the test contents, getting advice of subject matter experts or colleagues, preparing the test according to the plan, test administration and use, scoring the test and reviewing/analyzing test results. So the analysis showed that most teachers have shortage in using techniques help them to construct test items.
- According to finding obtained from interview conducted it was realized that writing exams plan before constructing exam items are not exercised. All the ten interviewees responded that they do not use any written plan personally and at department level. Planning the test which includes several activities that language teachers need to carry out before construct new exam such as draw up test 'blue print', specifying the content

,objectives of the test, type of items and time limit are seemed to be ignored by the teachers.

2. Concerning the principles that English language teachers should follow and do not follow:

- Most of the teacher did not use principles to be employed when they design classroom test such as encouraging their students to do a systematic review of materials they have learned through skimming the text book, skimming other exam materials and outlining major points and facilitating formation of study group before the test.
- Some of the teachers did not inform students to look quickly over the whole part of exam paper so as to get grasp of its different parts, remind students to mentally figure out how much time they will need for each part, advise students to concentrate as carefully as possible and alert students minutes before time is over so that they can proofread their answers during the test.
- Some of the teachers did not include feedback on specific items students did not do well except recording total result they scored on the answer sheet. Furthermore, they did not discuss the test questions with students and advise them to pay attention to points they did not do well in the future.

3. With regard to types of test items mostly used by the teachers:

- Most of the teachers use multiple choices item types repeatedly than the other types of test items in their exam. They usually use multiple choice items to relate their exams' item types with those of national exams and to get students to practice scoring better results in national exams. This was clearly seen from the data obtained from checklist analysis in the three consecutive semesters' teacher-made exams (2013 to 2015).

4. Concerning factors that affect teachers' practice in designing effective tests at classroom level:

- Majority of the teachers did not take and give training on test item designing practice at school level in their department. Furthermore, they have no experiences of saving exams in exam bank and storing in special area. Thus, there is limitation with these teachers in exchanging good practice of designing language tests among themselves.

Almost all of the teachers did not take any training regarding test designing practice at district education offices, regional education offices and ministry of education level to update their experiences of test designing skills except the one they took in measurement and evaluation course at higher institutions.

5.2. Conclusions

The following conclusions were made based on the above summary.

- Most of the teachers seemed to have shortage in using techniques suggested by scholars: such as planning the exam, writing exam objectives, outlining the test contents, decision and advice of subject matter (experts or colleagues), preparing the test according to the plan, test administration and use, scoring the test and reviewing or analyzing test results.
- Almost majority of the teachers did not use better test designing techniques such as using tests' written plan which includes several activities that language teachers need to carry out before constructing new exam. Specifying the content and objectives of the test, type of items and time limit seemed to be ignored by the teachers personally and at department level.
- Some of the teachers did not follow principles such as encouraging their students to do a systematic review of materials through skimming the text book, skimming other exam materials, outlining major points and facilitating formation of study group before the test, informing students to look quickly over the whole part of exam paper so as to get grasp of its different parts, reminding students to mentally figure out how much time they will need for each part, advising students to concentrate as carefully as possible and alerting students minutes before time is over so that they can proofread their answers during the test. Furthermore they did not include feedback on specific items students did not do well except recording total result on the answer sheet. Furthermore, they did not discuss test questions with students and advise them to pay attention to points they did not do well in the future.
- Most of the teachers used multiple choices item types repeatedly than the other types of test items. As the majority of the respondents attempted to provide their reason, they usually use multiple choice items to relate their exams' item types with those of national exams and to get students to practice scoring better results in national exams. This was clearly seen from the data obtained from checklist analysis in the three consecutive semesters' teacher-made exams (2013 to 2015).
- There are lack of attention to the problem by concerned bodies such as district education offices, regional education offices and ministry of education to give teachers training and update teachers' experiences of test designing skills.

5.3. Recommendations

Based on the above findings and conclusions, the following recommendations are forwarded.

- Since teachers are crucial agents to bring new changes on their students with regard to teaching and testing practice, they should update themselves with recent changes in the field of language testing and use better techniques to design effective tests. Therefore, the school should facilitate situations like budget, materials and time for department members to conduct in-staff training and to foster teachers understanding of test designing skills.
- Teachers should follow better principles when they design tests and implement at classroom level to meet their goal of teaching languages and try to help their students in order to make them better user of the language. Teachers should also help their students by revising tests questions with their students so that they can learn from their mistakes.
- Teachers should design as much as possible all types of test items to assess their students' language achievements rather than always prepare one type of test items (multiple choices). They should set short- answer item that is used to measure students' ability of interpreting diagrams, charts, graphs, and pictorial data. They should also use completion items to measure students' vocabulary knowledge in completing sentences and paragraphs.
- Concerned bodies in education sector like district education offices, zone education offices, regional education offices and the ministry of education should give focus to the problem and train teachers to update their awareness of designing effective tests since teaching and testing are two inseparable activities in the field of education to bring desirable changes on the students.

References

- Abedi, J. (2010). **Performance assessments for English language learners**. Stanford, CA: Stanford University, Stanford Center for Opportunity Policy in Education
- Alderson, J.C. and et al. (1995). **Language Test Construction and Evaluation**. Cambridge: Cambridge University Press.
- Asmare Mehret. (2008). “**An Assessment of Content validity of English Language Tests**”. (Unpublished M.A. Thesis). Addis Ababa University.
- Argaw Sileshi. (2004). “**The Content Validity of Grade 10 National English Language Examination**”. (Unpublished M.A. Thesis) Addis Ababa University.
- Arnold, H. (2005). **Language Testing**. Volume 22, Number 2, (page 126).
- Bachman, F. and Palmer, A. (1996). **Language Testing in Practice**. Oxford: Oxford University Press.
- Bachman, L. (1990). **Fundamental Consideration in Language Testing**. London: Oxford University Press.
- Berry, R. (2004). **The Research Project: How to write it, 5th edn**. London: Routledge
- Best, J. and Kahn, J. (1993). **Research in Education (7th.ed)**. New Delhi: Prentice Hall
- Brown, J.O. (1996). **Testing in Language Programs**. Prentice Hall Regents: Prentice Hall, INC.
- (2004). **Language Assessment: Principles and Classroom Practices**. Longman: Longman com.
- Cheng, L. (2005). **Changing Language Teaching Through Language Testing**. Cambridge: Cambridge University Press.
- Davies, A. (1999). **Dictionary of Language Testing**. Cambridge: Cambridge University Press.
- Dornei, Z. (2007). **Research Methods in Applied linguistics**. New York: Oxford University Press.
- Elbe, R. and Risible, D. (1991). **Essential of Educational Measurement**. Prentice-Hall of India: New Delhi.
- Fulcher, G. (2010). **Practical Language Testing**. London: Hodder Education, a Hachette UK Company.
- Gronlund, N. E. (1998). **Assessment of Student Achievement**. Boston: Allyn and Bacon. Last revised
- Harrison, D. (1983). **Language Testing Handbook**. New York: Cambridge

University Press.

Heaton, J.B. (1988). **Writing English Language Test**. New York: Longman Group.

-----, (1990). **Classroom Testing**. New York: London.

-----, (1991). **Writing English Language tests: A Practical Guide for Teachers of English as a Second or Foreign Language**. London: Longman Group UK Ltd.

Hicks, D. Little John, A. (1998). **Cambridge English for Schools (CES)**. Teacher's Book Level Two. Cambridge: Cambridge University Press.

Hughes, A. (1989). **Testing for Language Teachers**. Cambridge: Cambridge University Press.

-----, (2003). **Testing for Language Teachers**. (2nd ed.). Cambridge: Cambridge University Press.

John, P. (2004). **Writing and Reviewing Assessment Items: Guidelines and Tips**. University of North Carolina. Chapel Hill

Kifile Kebede. (1995). "An Assessment of Content Validity of High School English Language Tests". (Unpublished M.A. Thesis) Addis Ababa University.

Kopriva, R.J. (2008). **Improving Testing for English Language Learners**. New York: Routledge.

Koul, L. (1997). **Methodology of Educational Research**. New Delhi: Vikas Publishing House Pvt. Ltd. 1.

Kubiszyn, K., & Borich, G. (1984). **Educational testing and measurement: Classroom application and practice**. Glenview, IL: Scott, Foresman, pp. 53-55.

Lee, R. (2000). **Unobtrusive Methods in Social Research**. Buckingham: Open University Press

Linn, R. and Gronlund, N. (2000). **Measurement and Assessment in Teaching** (8th ed.) Columbus, OH: Merrill and Imprint of Prentice Hall.

Lyman, H.B. (1998). **Test Scores and What They Mean**. (6th ed). USA: Ally and Bacon.

Madsen, H. (1983). **Techniques in Testing**. Hong Kong: Oxford University Press.

Samuel Dossa. (2011). **"Assessing Content Validity of Primary Education Completion English Examination of SNNPRS"**: 2002 E.C in Focus. Unpublished MA Thesis. Addis Ababa University.

Selltize, C. (1959). **Research Methods in Social Relations**. New York: Holt, Rinehart and Winston.

Smith,H.(1975).**Strategies of Social Research**. Englewood Cliff, NJ: Prentice hall.

Yeshambel Almaw. (2011)“**An Assessment of the Content Validity of Grade 9 English Language Teacher-made Classroom Tests.**” (Unpublished M.A. Thesis)

Addis Ababa University.

**Adama Science and Technology University (ASTU)School of
Humanities and LawDepartment of English (Postgraduate
Program)**

Appendix A: Checklists for Analyzing Teacher-Made Test Items

Table 1: Multiple Choice Items analysis

		Exam Year			Re mar k
		2013	2014	2015	

No	Checklist Points	Yes	No	Yes	No	Yes	No	
	a. General Construction criteria of the questions							
1	The teacher distributed correct responses among the alternative positions.							
2	As, Bs, Cs, and Ds have nearly equal proportion as the correct answer.							
3	The teacher kept the number of alternatives at five and four.							
4	The teacher used negatively stated items in limited amount.							
5	The negatively stated items used were underlined and highlighted well.							
6	There was not more than one best answer to the stem.							
No	Checklist Points	Exam Year						Remark
		2013		2014		2015		
		Yes	No	Yes	No	Yes	No	
8	The teacher used direct question rather than included as part of the stem, so that all options can be clearly distinguished							
9	The teacher included much of the item in the stem and kept the alternatives short and brief.							

.	C. Construction of the options of the questions							
10	The teacher kept all the alternatives in a similar format. (i.e. all phrases, all words, all sentences.)							
11	The teacher listed options on separate lines rather than included as part of the stem so that all options can be clearly distinguished.							
12	All options were plausible responses to the stem?							
13	The teacher didn't include poor alternatives for the sake of having more options.							
14.	All of the options had grammatical consistency							
No	Checklist Points	Exam Year						Remark
		2013		2014		2015		
		Yes	No	Yes	No	Yes	No	
	with the stem (tense, number, etc...)							
15.	All the alternatives were written approximately in the same length.							
16.	Capital letters (A. B. C. D.E) were used as response signs rather than lower case letters.							

Source: - Adapted and formulated into checklists from:

_____.2012/13.**Language Assessment, Evaluation and Testing (TEFL623)**: Adama: ASTU: Ethiopia.

_____.2006.**Item-writing, Guidelines and Suggestion** as derived from 46 authoritative text books: _____:

Table 2: Analysis of Teachers' Techniques in Selecting Reading Passages

No	Checklist Points	Exam Years						Remark
		2013		2014		2015		
		Yes	No	Yes	No	Yes	No	
1.	The passage has a beginning, middle and end.							
2.	The passage is written well in standard English containing facts.							
3.	The passage was come from other sources apart from text book.							
4.	The passage could be interesting and to the level of students target grade level.							
5.	The passage’s length is the appropriate to the grade level							

	(300-1200 words).							
6.	The teacher presented the passage that includes students from variety of cultural back grounds.							
7.	The passage’s content was free from unnecessary thoughts that burden works on students.							
8.	The teacher wrote the passage that advocate common value to all students taking the test.							
9	The passage is free from problems of bias or sensitivity (i.e. avoid words							
No	Checklist Points	Exam Years						Remark
		2013		2014		2015		
		Yes	No	Yes	No	Yes	No	
10	The teacher sited the source of the passage. (Author’s name, year of publication, place of publication...etc)							
11	The passage is free from over using idiomatic expressions or figurative language unnecessarily to students whose first language is not English.							
12	The teacher avoided topic in his passage that is likely to upset, or unduly distract students and affect their performance on the test.							

Source: - Adapted and formulated into checklists from:

_____.2012/13.**Language Assessment, Evaluation and Testing (TEFL623)**: Adama: ASTU:
Ethiopia.

_____.2006.**Item-writing, Guidelines and Suggestion as derived from 46 authoritative text
books: _____:**_____

Alderson, J.C. et al.1995.**Language Test Construction and Evaluation**

Adama Science and Technology University (ASTU)
School of Humanities and Law
Department of English (Postgraduate Program)

Appendix: B

Semi-structured Interview items for Teachers

Dear Teachers,

At present, I am doing my research work on teachers' practice of designing English language tests at classroom level in grade ten. The following interview questions are designed to collect relevant data for the study. Thus, your responses will have much contribution to the success of the intended study. You are therefore kindly requested to listen to each item carefully and give your genuine responses. Concerning the information you give me, I would like to assure you that all would be kept confidential and used only for the research purpose.

Thank you!

1. How often do you give final exams for your students in a year? Do you take part in all final examinations set at department level?
2. Why do you give final examination for your students?
3. Do you think that your students can be benefited from the examination you give them? If say "yes", how they can be benefited from such examinations?
4. As a group (English department members) or as individual, are there techniques you always use to construct examination items? Could you tell me some of these techniques and the way you use them?
5. Do you have written plan to give examination to your students? If you have written plan, what activities do you include in the plan when you write your plan?
6. In your perception, what are the uses of these activities such as writing objectives of the examination, outlining details of examination content?
7. What are the pre activities that you do before you distribute exam papers for your students and after you score students' answer sheets?

8. What factors do you think can affect students' psychological emotions during the time they take exams at class room?
9. Do you think that your students exactly know content area your exams covers? If you say 'yes', how can they know it?
10. How can you bring exam languages to the level of your students' language proficiency and to their real life situation?
11. After you scored students' answer sheet do you return back to test items and review with the students? What is the importance of such works?
12. Which type of test items do you most of the time use in your test? And why do you use it ?
13. In your department do you run training program on test designing practice and have you exam bank and other special exam reserving store?
14. Have you ever taken any training on test designing conducted by ministry of education, regional education office and district education office sometime at somewhere?

Appendix C: Samples of teacher- made exams.

- **See them. They are attached at the end of this thesis paper.**

Appendix D: Interviewees responses on each interview questions

A. Interviewees' responses regarding their awareness about exam purposes

Question1: How often do you give final exams for your students in a year? Do you take part in all exams set at department level?

Teacher 1: We give twice in a year. Yet, I haven't take part in all exams set at department level. But I remember I took part in preparing final exams of 1st semester from 2013-2015.

Teacher 2: We actually give final exams twice per a year. We always prepare only the first semester final exam at department level. But the second semester final exam is prepared at national level as model. So, we don't prepare the second semester final exam at school level.

Teacher 3: We give final exam two times in a year and each semester one. Yes, I take part in designing exam items in the department.

Teacher 4: We give final exam twice in a year. As, I am one of the department members, I always participate in constructing exam items at department level.

Teacher 5: We only prepare the 1st semester. This is because of that, the second semester final exam is prepared at national level and comes from there.

Teacher 6: We prepare only the first semester. We do this, because the second semester final exam is set at national level as model. So, we don't prepare two times we want to save our time.

Teacher 7: We design only the first semester exam. Most of the time, the 1st one is prepared at national level.

Teacher 8: We prepare only one the first semester. But we give final exam for our students twice in a year together with that of model exam.

Teacher 9: We give two times within a year. But as a department, we always prepare only the first semester one.

Teacher 10: We prepare only one exam, that is the first semester, but we don't prepare the second semester one.

As it can be understood from the above interviewee's responses, all of the interviewees responded that they give final exams to grade for ten students twice within a year. Of the two times exams, they take part in designing the exam items only in the first semester one. As it is understood from the reason they gave, they prepare the first semester only. They do not prepare

the second semester final exam. They use the model exam that is prepared at national level instead of second semester final exam.

Question 2: Why do you give final exams for your students?

Teacher 1: I give test (exam) when I have special reason. My special reason to give my students' exam is to evaluate their understanding of the subject I taught them. But I don't have any reason to give my students exam.

Teacher 2: I give final exam based on my need. When I want to assess students' total achievement about the subject matter they had already learnt.

Teacher 3: I give final exams to check students' problems and to enable them study hard in the future. I rely on my need.

Teacher 4: I give final exam to know whether or not the students mastered the main three language skills (reading, writing and speaking). I also give them to know whether or not all students understood the materials taught in the text book. So, I based on my reason.

Teacher 5: I give exams to check what they gained and lost from what they had learnt. I don't have special reason to give tests.

Teacher 6: I give exams to get my students prepared for the national exams. This is my reason.

Teacher 7: I give exam to evaluate my students' success and failure on what they had already learnt. And to provide them offer like summarizing, make up etc. So I have my own reason to give exam.

Teacher 8: I give exam for my students to check their weakness and strength regarding what they had already learnt.

Teacher 9: I have my own reason to give exam for my students. The first reason is to measure students' understanding. The second reason is to promote students from grade to grade.

Teacher 10: To check whether or not students are on the right way of learning. Thus I give exams for my students depend on my own need.

As it can be analyzed from the above responses given by all the interviewees, most of the interviewees said that they have reasons to give exam to their students. Almost their reasons to give exams for their students revolve around evaluating what students had already learnt.

Question3: Do you think that your students can be benefited from the examinations you give them? If you say ‘yes’, how can they be benefited from such exams?

Teacher 1: Yes absolutely. They can be benefited from the exam they take in the class room. The exam helps students to evaluate themselves. It tells them where their difficulty is in their study. It helps them as tool to indicate their difficult areas.

Teacher 2: Yes, it helps them. The exam enables students to be informed what they gain or loss. It also encourages them to give focus where they must do more in their study.

Teacher 3: yes, it helps students to evaluate their academic achievements. It is used as indicator to locate students’ difficulty area.

Teacher 4: Yes it helps my students. My exam can show them the way how they should get prepared for the national exams.

Teacher 5: Yes, it can help them and students can be good beneficiary from my exam. My exam puts my students to go on the right way of learning.

Teacher 6: Yes, students can be benefitted from the exam I give them. The exam I give them can be used as best tool to indicate students to study in accordance with national exams.

Teacher 7: Yes, it can be help students’ exam can encourage students to study more in the future.

Teacher 8: Yes, exactly. Students can be benefitted from the exam they take. They get motivation, it enables them know their weakness in their lessons.

Teacher 9: In my belief, students can be benefitted from my exam. They are encouraged to study more for the next exam.

Teacher 10: Yes, students can be benefited from what they invigilated. It enables students to learn from their mistakes.

According to the responses given to the third question, all of the interviewees believe in about the importance of giving exam for students. All the interviewees responded that their exams have great importance for their students. They reported that they give their students exam to evaluate students academic achievements and students can be benefited from the exam: by evaluating their weakness and strength on the content area, by looking back to the content area where they have difficulties, and by being motivated to study more. But few of the interviewees responded that the exam helps students to show the way national exam is prepared and to encourage them study more to promote from grade to grade. Speaking generally, it seems that most teachers have concept about the merits of their exam to their students.

B. Interviewees' responses regarding their experiences in using test techniques that help them to design exam items.

Question 4: As a group (English department members) or as individual, are there techniques you always use to construct exam items? Could you tell me some of these techniques and the way you use them?

Teacher 1: Yes, we have common techniques that we use to construct exam items. Some of them are first; we discuss and decide the type of exam items we are going to include in our exam. Next, we share types of items each other. Finally, each teacher prepares one item the items according to their logical order. At the end we evaluate the question being with exam committee.

Teacher 2: Yes, we have. We decide the total number of items, types of items and share the questions to construct personally; finally, we bring what have constructed together.

Teacher 3: Yes, a department, we have common techniques to develop exam items. First, as a group member we come together and decide the type of exam items and total number of items. Then, we bring together what have prepared and decide on each item.

Teacher 4: we don't have special techniques, we prepare exam at department level after we reached agreement.

Teacher 5: we prepare exams at department level. We organized exam items when we reach an agreement. After that, we check each items difficulty level whether or not they match or miss-match our students name.

Teacher 6: I don't have special techniques personally, but, we have common techniques as department, that is we get gathered together to share test items type in the department. Each teachers is given one type of item and evaluation when evaluation over. We get printed of test and deliver to our students.

Teacher 7: I don't have special techniques. What I remember is that I learned some techniques when I was in higher institution. But I don't exercise it now.

Teacher 8: Yes, we have common techniques. One of our common techniques as department member is to include all language skills in the exam items.

Teacher 9: yes, I have my Owen techniques. I always refer to the previous exams to get some information.

Teacher 10: I don't have my own techniques, but we have common techniques as English department members. We always decide to include language skills together in the exam.

Mainly the above responses presented by the interviewees almost inclined to be techniques that they use in designing exam items at department level. Similarly all of the respondents declared that few item types are shared for teachers. Then, a teacher constructs types of items he or she got. At end they come to the department and decide on the exam items. On other hand, few of them could tell some effective techniques suggested by scholars, such as planning the exam, writing exam objectives, outlining the test contents, decision and advice of subject matter experts or colleagues, preparing the test according to the plan. Test administration and use, scoring the test and reviewing/analyzing test results.

Question 5: Do you have written plan to give exams for your students? If you have written plan, what activities do you include when you write you plan?

Teacher 1: I do not use any written plan to prepare the test or any exam. I think, it is not must to have written plan.

Teacher 2: I do not have any written plan personally. Similarly I have not seen any one in our department who uses written plan to give tests.

Teacher 3: It may be depend up on teachers need. As to me I do not have any written plan and I do not use it. Because, I am not familiar with this type of technique, I do not use it.

Teacher 4: we do not have any written plan. I have never seen anyone using written plan to design test/exam in my department.

Teacher 5: I do not write plan to construct exam items.

Teacher 6: I have no any plan concerning exam item writing process. So I, do not use it.

Teacher 7: To tell you the truth, we do not have any written plan. And no one prepares his/ her test using it.

Teacher 8: We do not have any written plan. We only arrange schedule when give tests/ exams at department level.

Teacher 9: We have plan of giving exam ideally but not in written form.

Teacher 10: I do not have any written plan. We have schedule at department level. We only use it.

From the above interviewee's responses it can be realized that writing an exam plan before constructing exam items is not exercised. All the ten interviewees responded that they do not use any written plan personally and at department level. Planning the test includes several activities that language teachers need to do before constructing anew exam. Some of these activities are drawing up test 'blue print, specifying the content and objectives of the test, type of items and time limit. As it can be understood from the interviewee's responses they do not have practice of writing test plan.

Question 6: In your opinion, what are the uses of such activities like writing the objectives of the examination and outlining details of examination contents?

Teacher 1: Since I do not use them I do not know their importance so much. But, I assume they seem important if they are practiced.

Teacher 2: I believe, they are very important though we do not use them in our school.

Teacher 3: I do not know their use in details, but it is good if they are practiced well.

Teacher 4: We do not use any of them. So, I cannot tell you their use.

Teacher 5: I belief that they have importance. They direct teachers as to how they prepare exam items, they help teachers to sample all learned contents etc.

Teacher 6: In my point of view, they help teachers to prepare good exam. But we do not have such practice of using these things.

Teacher 7: Yes, of course, these activities can have their importance. They can guide teachers to construct test items.

Teacher 8: I believe that these activities can indicate teachers why to give exams, when to give exams where to give exams and to whom exam is given.

Teacher 9: Though, I do not use them. I belief they can be very important to write exam items.

Teacher 10: Yes. They have their own uses they keep the teacher not to go beyond their ability.

As it can be observed from the response above, most of the interviewees believe in the importance of using activities such as writing objectives of the exam, outlining exam contents for teachers. However, none of them could tell the practical implementation of these techniques in their school.

Question 7: What are the activities that you do before you distribute exam papers for your students and after you score students answer sheet?

Teacher 1: The pre activities that I do before distribute exam papers are: arranging sitting rows, avoid crowded sitting. Selecting better students to the front sitting and I tell them not cheat. I do nothing except giving students their answer sheet.

Teacher 2: I check whether or not the question papers have all items, I select better students to the front part. After, I score; I give students their answer sheet.

Teacher 3: Before I get my students start the exam, I arrange their setting. I check the question and I distribute exam paper for it. After, I score students' answer sheet, I record the result and give students their answer sheet.

Teacher 4: Before I give students the question paper, I arrange benches, tables and check any written thing on the black board. After I score I revise the question with students.

Teacher 5: I check total exam items and distribute the question papers for students. After, I score the students result, I realize the type of items and the number of items most of the time lost and I give tutorial on that content area.

Teacher 6: I tell students to write their names and fill necessary information. I also check test items, No, of questions, Page numbers and clarify of questions. After I score the answer sheet I see the item most of the students lost and try to give clues on it.

Teacher 7: Before the exam, I tell them to write their name and information. I encourage students to tell me the items they find difficult I do not do except recording their results.

Teacher 8: I tell them the instruction, reason of that test and check question. After I score I record students result and give them their answer sheets.

Teacher 9: Before the exam, I tell them not see their text book and exercise book not to copy from their friends. After I score I record students result and give them their answer sheets.

Teacher 10: Before the exam is started, I check questions and make necessary correction and I see the sitting arrangement. After, I score the test, I write answers key. Next, I use the answer key and score. Finally, I record the point and give students their answer sheet.

It is possible to generalize that most respondents listed activities that they do before distributing the exam papers for their students. As their responses reveal, they arrange sitting, avoid noise, check questions on all pages, tell students to fill necessary information, look at each instruction and to concentrate as carefully as possible on question and alert students mentally how much

time they use for each item type. On other hand, important activities to be done like revising the exams with students, writing specific feedback on questions a student lost answers and areas to be focused by students in their future were not mentioned largely.

Question 8: what factors do you think can affect students' psychological emotion when they take exams in the classroom?

Teacher 1: In my opinion, factors that can affect students' psychological emotions during they take time are; students readiness to take exam, invigilators personal behavior, noise sounds and types of food they eat the exam day.

Teacher 2: I think, students always consider English exam more difficult than other exams. This can affect their psychological emotions.

Teacher 3: Students' psychological emotions can be affected by different conditions, such as coming late to the exam room, illegibility of written material and teachers' personal behavior itself.

Teacher 4: Some unnecessary things that happened in the class disturb students not do well. They may be sitting arrangement unclear instruction and noise sounds around the school.

Teacher 5: The difficult words that exist in the exam back ground of students family such as their economic status, peace among the family members and the culture of the students can be factors those affect students psychological motion when they take exams.

Teacher 6: The setting by itself can bring impact on students' psychological emotion. When I say setting I mean the time, the place where and when students take place can affect students' psychological emotions.

Teacher 7: I think fear of exam can affect students' psychological emotion. I do not think, the rest can affect more than it.

Teacher 8: the clarity of instruction and items by itself can have such problems on students' emotion.

Teacher 10: Teachers way of constructing tests can affect students' emotion, in addition to that the speed that other higher achiever students use may bring negative impact on medium and lower achievers. If the high achievers students do at high speed, other may feel fear and become worried.

Noticeably, the majority of the interviewees were inclined to answer some factors that can affect students' psychological emotion when they take their exam in the classroom. As their responses reveal, students' emotions can be affected in the classroom by invigilators' behavior, speed their peers use on doing exam, fear of exam, vague exam instructions and students' personal problems.

**Question 9: Do you think that your students exactly know content area your exam covers?
If you say "yes" how they can know it?**

Teacher 1: I think that except clever students those study hard, the rest do not know where the test area covers. If they listen carefully when I tell them the area exam cover they can know it.

Teacher 2: I do not think that all students know the content area of the exam. Few of them can know the areas. If they attentively listen to the advice I give them they can know it.

Teacher 3: Few of the students may know the content area of the exam. This because of that most students do not give attention to the advice their teacher give them.

Teacher 4: I think they can know it. The reason why I could say that is I always tell them the beginning and ending pages of the text book from which the exam will be set out.

Teacher 5: I do not all can understand and know the content area that is covered by the exam. I always tell them only the unit of the lesson where exam is come from.

Teacher 6: In my view this is impossible for students to know the content area. As my exam is language exam, it can include different contents.

Teacher 7: Some can do that but not all students do those things. Better students study carefully. As a result they can know the area

Teacher 8: I am not sure enough that students know the content area of the exam. In my part, I tell them the wait, type of items and total number of items.

Teacher 9: If I revise the content area, they may it, unless I do that, they cannot know it exactly.

Teacher 10: It depends on the students' ability. If they are clever enough, they know the area through depth study: otherwise they cannot know it.

Out of the respondents who were interviewed, majority of them appeared to give relatively the same responses. They are not sure enough whether their students exactly know the content areas exams cover and only few of them help their students know the content areas of the exams.

Question 10. How can you bring exam language to the level of your students' language proficiency and to their real life situation?

Teacher 1: I try to use simple language as much as possible and try to use words students know well.

Teacher 2: I use words contextually rather than use them as they are. In this way, I try to simplify language difficulty for my students.

Teacher 3: I try to change some foreigners name into our local names students may know. This is one of my techniques to decrease language difficulty.

Teacher 4: If I come across with difficulty words in the passage, I delete them and use another simple word instead of them.

Teacher 5: In my view, it is better to select simple words to use in the exam. When words become simple in a sentence, students can understand the message.

Teacher 6: To make language easy to my students in their exam, I use my dictionary and select easy words.

Teacher 7: If words are too difficult for students in the exam, students are put in the anxiety. So, to reduce the anxiety, I try to use difficult words in context.

Teacher 8: To make my exam language easier, I select words according to students' level and use them I take care not to use difficult words.

Teacher 9: I try to adapt the language of exam to the students' culture and real life. I try to substitute difficult words with their easy balanced words.

Teacher 10: I try to use words those I think familiar with students. I try to relate them to their cultural background and real life situation.

Except few of the interviewees, most of them said they take care of the language difficulty when they prepare exams for their students. They usually use simple languages to construct exam items.

Question 11: After you scored students' answer sheet, do you return to test items and revise them with your students? What is the importance of such practices?

Teacher 1: To tell you the truth, I do not do this always. But when I found the students scored below pass mark. I return back and discuss the question with the students.

Teacher 2: I do not do it. But I believe it helps both students and the teacher. The reason why I could say that is students can learn from their mistake and teachers need to think on his teaching methods.

Teachers 3: Yes. I return back and revise the questions with my students. It helps students to know the correct answer.

Teacher 4: I return to the exam and discuss it with students. It helps students to check their mistakes.

Teacher 5: I do not do that. I only revise the question with my students when I found their results below my expectation

Teacher 6: I do not revise it most of the time since it is important is not clear for me. May be, when I find it is important to do that, I revise it so more times; but I do not do all the time.

Teacher 7: I revise it. When I revise the exam students are motivated and study hard for their future exams.

Teacher 8: Most of the time, I write correct answers on chalk boards and tell students to get copy of it.

Teacher 9: I do not have such experience. Rather I give them make up and tutorial class.

Teacher 10: In my part, I do not return back again to the exam the students have already taken. I believe this is wasting time. Instead of doing that it is better to teach them contents related to their exam before they take the exam.

Most of the responses that were given by the interviewees seem to have less relevance to some of the suggestions provided in the review of literature part of this study. Generally, most of the teachers have no clear reasons to revise their exam items with their students.

D. Teachers' responses regarding types of test items they use or do not use most of the time in their exam.

Question 12: Which types of test items do you use most of the time? Why? Which ones you do not use most of the time? Why not?

Teacher 1: as a department and as individuals, we always use multiple choices. We use it to get our students use multiple choice to study wide area of contents and we use it to get our students in order to exercise national exam. We use another type in continuous assessments.

Teacher 2: I use multiple choices. Multiple choices can touch several area of the lesson content. Though I use it, I do not believe it can foster students' creativity such as writing paragraph and essay. As a habit we always prepare multiple choice items. But most of the time, we do not use other types.

Teacher 3: I use multiple choices. I think it helps me to write conversation, jumbled sentence and vocabulary and others. Other types of test items cannot do what multiple choices can do. Due to this reason, we do not use others in most cases.

Teacher 4: In our school, when we prepare exams for our students we use multiple choice items types. We depend on the grade ten national exams and prepare our exam accordingly. Because of, we do not see other types of test items in national exams except the multiple choices, we do not use them.

Teacher 5: We use the multiple choices. We use it because it can include all types of items. Most of the time, we do not use other types of test items like true or false, completion and essay in grade ten final exams. These type of item types do not appear in national exams.

Teacher 6: I use multiple choices. It is convenient to write. It also saves our time of constructing items. In addition to its easiness to prepare, it gives teachers to assess large portion of the contents. When we think others, they have no power to this.

Teacher 7: We use multiple choices. We write this type of items in our department. To tell you the truth, I do not know the reason.

Teacher 8: We use multiple choices. We have habit of writing it. But, when I think about it is importance, I did not find it better than other item types. This is to say that it does not enable students to show their creativity power.

Teacher 9: We use multiple choices. It helps our students to assess most of their lesson contents. It is also helpful for teacher to construct immediately. We do not use most of the time other item types because, they do not do these activities.

Teacher 10: We use multiple choices especially in our final exams. We do this in order to include many samples of language contents in our exam. Again another reason to use multiple choice is to provide our students a chance of remembering what they have learned.

As it could be understand from the responses given by the interviewees, most of the teachers seem to use multiple choice item types repeatedly than the other types of test items. The responses provided by majority of the interviewees tend to show the primarily reason for teachers

to use multiple choice is to relate teacher-made item types with those of national exams and to get students practice how to answer similar exam items appeared in national exams, to include many sample of language contents in the exam, and it is easy to score.

E. Responses given by teachers regarding factors that can affect their test designing practices

Question 13: In your department, do you run training program on test designing practice? Do you have an exam bank or any other special exam reserving or storing mechanism?

Teacher 1: In our department, yet we have not conducted any training regarding methods and techniques of exams preparation. We do not have exam bank and special area of exams store. If we want to put it, we can put it in the library.

Teacher 2: still, I have not taken any training on exam preparation in our department. In the same way, we do not have exam bank and special exam storing places.

Teacher 4: I did not take part in any training of exam preparation. Since I, came to this school) I have never seen such kind of training in our school.

Teacher 5: I think it is good if this type of training is given in the school. Unfortunately we did not conduct the training and we did not have special area to store exam and exam bank in our school.

Teacher 6: I have not seen any test designing training conducted as staff in our school. We do not have exam banks and special exam storing places. I do not know the reason.

Teacher 7: we did not conduct training in our department regarding how to design exams for our students'. In the same way, we do not have exam bank as most of us have no computer's skills. Similarly we do not have special area to save our exams.

Teacher 8: Totally, we did not conduct test designing training and we do not have any exam bank and special storing places.

Teacher 9: We did not have such experiences. We do not put exams in special area. Even I do not know the term exam bank.

Teacher 10: we do not conduct training on test designing; instead we share each other, our experience of test designing orally. The other thing, you have asked me about exam bank and place to store exam. I tell you the fact; we do not have such experience of keeping our exams in special places.

Regarding the question about their practice in running training on test designing techniques, majority of them reported that they have not taken training on test item designing practice at in their department and have no experiences of saving exams in exam bank and storing in special area. This shows that there is limitation with these teachers in exchanging good practice of designing language tests among themselves.

Question 14: Have you ever taken any training on test designing conducted by the ministry of education, regional education office and district education office?

Teacher 1: No, I have never taken it at any level. I have not gotten such chances.

Teacher 2: I have not taken it at any level. I have not heard about test designing training.

Teacher 3: I did not attend test designing training. I have been teaching English since 2002 in the high school, but yet I have not gotten that chance.

Teacher 4: I did not take it anywhere. I do not know the reason.

Teacher 5: Absolutely, I did not take training on test designing at woreda level, Oromia or ministry of education. But what I do not deny is that I took this course training when I was in higher institution.

Teacher 6: I remember, I took the course called measurement and evaluation at my BA level. I learned this type of training on test designing.

Teacher 7: I did not take elsewhere the training except the one I had taken in the university.

Teacher 8: I did not take elsewhere the training. That is to say, at District level regional and national level. But, I studied it in one of my courses during. I was studying in the university.

Teacher 9: There is no training on test construction. As to me, I took training on methods of teaching in the university. Except this one, I did not take training test practices anywhere.

Teacher 10: I have never taken this kind of training. What I want to clarify for you is, I learned about test in measurement and evaluation courses in higher institution. After that I did not take part in training of test designing at woreda, region and national level.

Almost all of the interviewees responded that they have never taken any training regarding test designing practice at district education offices, regional education offices and the ministry of education level to renew their experiences of test designing skills except the training some of them took in measurement and evaluation course at higher institutions.

