

**ADDIS ABABA UNIVERSITY
INSTITUTE OF LANGUAGE STUDIES
DEPARTMENT OF FOREIGN LANGUAGES AND
LITERATURE
(GRADUATE PROGRAM)**

**AN ASSESSMENT OF THE CONTENT VALIDITY OF
ENGLISH LANGUAGE TESTS: THE CASE OF AWASSA
COLLEGE OF HEALTH SCIENCES**

**BY:
ASMARE MEHRET**

**JUNE, 2008
ADDIS ABABA**

**AN ASSESSMENT OF THE CONTENT VALIDITY OF
ENGLISH LANGUAGE TESTS: THE CASE OF AWASSA
COLLEGE OF HEALTH SCIENCES**

**By:
Asmare Mehret**

**A Thesis Presented To the Department of
Foreign Languages and Literature
(Graduate Program)**

**In Partial Fulfillment of the Requirements for
the Degree of Master of Arts in TEFL**

June, 2008

**ADDIS ABABA UNIVERSITY
INSTITUTE OF LANGUAGE STUDIES
DEPARTMENT OF FOREIGN LANGUAGES
AND LITERATURE
(GRADUATE PROGRAM)**

**An Assessment of the Content Validity of
English Language Tests: The Case of
Awassa College of Health Sciences**

By:

Asmare Mehret

Approved by Examining Board:

Advisor

Examiner

Examiner

Signature

Signature

Signature

ACKNOWLEDGEMENTS

My deepest gratitude goes to my advisor, Dr. Nuru Mohammed for his regular guidance and constructive advice which have enabled me to organize this thesis in the present form. I would also like to extend my thanks to Ato Endrias Yishak and Ato Abay Banti for they helped me in different work aspects. Moreover, I would like to express my heart felt thanks to w/o Mulualet Workneh for her typing the material, and Addis Ababa University for financial assistance it provided.

TABLE OF CONTENTS

	PAGE
Acknowledgement	I
List of Table	iv
List of Appendices	v
Abstract	vii
Chapter One	
1. Introduction.....	1
1.1. Background	1
1.2. Statement of the Problem	3
1.3. Objectives of the Study	4
1.4. Research Questions	4
1.5. Significance of the Study	4
1.6. Scope of the Study.....	5
1.7. Limitation of the Study	5
Chapter Two	
2. Review of Related Literature	6
2.1. What Is Test?	6
2.2. The Link between Teaching and Testing	6
2.3. Types of Tests	9
2.4. Continuous Assessment	12
2.5. Qualities of a Good Test	14
2.5.1. Practicality	14
2.5.2. Reliability	15
2.5.3. Validity	16
2.5.3.1. Face Validity	16
2.5.3.2. Concurrent Validity	17
2.5.3.3. Predictive Validity.....	18
2.5.3.4. Content Validity	18

2.5.3.4.1. Importance of Content Validity to Language Learning.....	20
2.5.3.4.2. Guidelines to Establish Content Validity.....	21

Chapter Three

3. Methodology.....	23
3.1. Data Sources	23
3.2. Sampling Materials	23
3.3. Data Collecting Procedures.....	24
3.3.1. Course Book Analysis.....	24
3.3.2. Interview to Teachers	24
3.3.2. Sample Test Analysis.....	25

Chapter Four

4. Results and Discussion	27
4.1. A Brief Overview of the Course Book.....	27
4.2. Analysis of Content Areas of the Course Book	28
4.3. Analysis of Sample Test Papers	30
4.4. Discussion	36

Chapter Five

5. Summary, Conclusion and Recommendations	40
5.1. Summary	40
5.2. Conclusions	41
5.3. Recommendations.....	42
Bibliography	43
Appendix	45

LIST OF TABLES

	Page
Table 1: content area of the course book and frequency of items of practice by content areas.....	26
Table 2: Summary of content of the coursebook areas on the basis of period allocated	28
Table 3: – Summary of total frequency of items in various content areas in sample test papers by content areas and test year.....	29
Table 4: Total frequencies of content areas of course book and sample test papers	32
Table 5– The degree of relationship between the course book and sample test papers	33

LIST OF APPENDICES

TITLES	Page
Appendix 1A: Frequencies of content areas of test one in 2006 by question number and content area	42
Appendix 1B: Frequencies of content areas of final test in 2006 by question number and content area	43
Appendix 2A: Frequencies of content areas of test one in 2007 by question number and content area	44
Appendix 2B: Frequencies of content areas of final test in 2007 by question number and content area	45
Appendix 3A: Frequencies of content areas of test one in 2008 by question number and content area	46
Appendix 3B: Frequencies of content areas of the final test in 2008 by question number and content area	47
Appendix 4A Test one: Awassa College of Health Sciences English test one for third year students -2006	48
Appendix 4B Test two: Awassa College of Health Sciences English final examination for third year students 2006	51

Appendix 5A Test one: Awassa College of Health Sciences English

test one for third year students -2007..... 57

Appendix 5B Test two: Awassa College of Health Sciences

English final examination for third year students -2007 60

Appendix 6A Test one: Awassa College of Health Sciences English

test one for third year students -2008..... 67

Appendix 6B Test one: Awassa College of Health Sciences English

final examination for third year students -2008..... 70

ABSTRACT

The purpose of this study was to assess if the coverage of English language tests administered in Awassa College of Health Sciences (ACHS) fairly represent the coverage of the coursebook. Hence, to meet the purpose of the study, English tests for third year students were selected. The main data sources, therefore, were the English coursebook for third year students and sample test papers administered for the same group of students in the academic years 2006-2008. From each year, one progress test and one final classroom achievement tests were included in the study. Then, the contents of the coursebook and sample test papers were analyzed in order to get the required data for the study. Besides, unstructured interview was administered to all English language teachers in the college.

After the necessary data had been obtained, Siegel's formula was used to determine the extent to which the two observations (Coursebook and sample test contents) are related. The result of the study shows that the content of sample test papers does not adequately represent the coverage of the coursebook. Some content areas, for example, grammar pronunciation, listening in the coursebook were not proportionally incorporated in the sample test papers. That is, there was a problem in allocating the percentages in tests from each content area in the teaching material coursebook. The degree of relationship between the coursebook and sample test papers was found to be 0.13. Thus, it was concluded that English language tests in Awassa College of Health Sciences were poor at content validity since they were unrepresentative.

CHAPTER ONE

1. INTRODUCTION

1.1 Background

Many scholars, for example, Bachman (1990); Alderson (1991); Weir (1993); Teshome (1995) and others agree that testing plays a considerable role in teaching- learning process. Weir (1993), for example, notes that testing aids teaching learning in that it helps teachers encourage their students towards learning, assists teachers to monitor their students' progress in learning etc. Moreover, feed back from tests provide information to the concerned bodies to revise a program or curriculum. In line with this idea, Guthrie (1984) in Nuru (1992) explains "Testing is assumed to play a considerable role in such important matters as the development of curriculum objectives as well as the production of textbooks and other instructional materials." Similarly, weir (1993) says, "A test contributes to the improvement of curriculum." In a general sense, we can say a test is another continuum of teaching, more importantly, in formal educational situations.

However, to offer such supports to teaching learning process, a test has to satisfy three major criteria of a good test. According to Hughes(1989) and Weir(1990), a test should be valid – test what is intended to be tested, reliable scores should be consistent when repeated and practical – a test should be economical and easy to administer and score. As a result, scores from such tests can convey reliable information to teachers, testees, and material developers.

Though validity, reliability and practicality have a mutual contribution to the good quality of a test, this study focuses on validity. More specifically, to content validity of classroom (teacher made) tests. It focuses on progress and achievement tests. This is because content validity is the most important

aspect of validity. Similarly, progress and achievement tests are more influential tests in formal educational setting; in addition, they are tests which nearly all teachers have an access to prepare and are. So, the researcher feels that a study on content validity, especially, in relation to classroom tests is worthy.

Content validity, refers to the concept, “whether the test produces a reasonable sample of the contents of the syllabus,” Underhill (1991). As it was said earlier, content validity is more important than other validity aspects such as face validity, Concurrent validity and predictive validity. Hughes (1989) says, “Content validity is quite relevant because it is a means to ascertain how thoroughly a language tests sample the instructional objectives or the universe of a certain behavior.” It helps to check student’s performance or level of progress in each content areas of the syllabus. In other words, it assesses every objectives of the course attended by the students. Likewise, Bachman (1990); Heaton (1988 and 1990); Weir (1990) regard content validity as an important feature of tests.

Thus, from concepts said about content validity, one can learn that test constructors in general, language test writers in particular, should strive to establish content validity of their tests. Tests should properly sample from the objectives of a given course because such tests give an opportunity to students to demonstrate their proficiency of language in each content areas of the syllabus. Consequently, statements that can be drawn by teachers and other concerned authorities concerning students’ level of performance can be meaningful. More importantly, tests help students to identify in which course content areas they are weak or/ and strong at. To conclude, testing scholars like Weir(1993) say contents of classroom tests must reflect what has been in the syllabus and treated in the classroom to arrive at relevant decisions about students, syllabus etc.

In local context, some studies have been conducted on testing areas. Studies by Teshome (1995); Kifle Kebede (1995); Nuru Mohammed (1992); Tibebe Alemayehu (1992) and Alemu Tsegaw (1983) are examples. Some of these studies were conducted on validity aspect. For example, Tibebe (1992) has conducted a study on the predictive power of ESLCE English in 1987/88 compared to two locally designed integrative language proficiency tests. He has found that the predictive validity of ESLCE in 1987/88 was found to be weak. Kifle (1995) has also conducted a study on content validity of grade ten English language tests with reference to the former textbook (English for New Ethiopia) which was structural approach based. Alemu (1983) has also conducted a research on National Examinations for grades six and eight. One of the specific objectives of the study was to see the content validity of the examinations. The result of the study shows that the examinations lack content validity.

The present study also focuses on assessing content validity of English language tests at college level. It aims to search for the extent to which English language tests at Awassa College of Health Sciences are valid contentwise. As we have seen studies by Kifle and Alemu, content validity of each sample tests is found to be weak. How about in advanced levels? Are tests in colleges, where various pre-conditions (example, class size) that guarantee to the goodness of tests are supposed to be in a better condition, good at content validity or not? This study will answer the question.

1.2. Statement of the Problem

Test quality can be affected by a number of factors. One of these factors is the adequacy of sampling test contents from a defined syllabus. A test should sample a wide range of important language contents discussed in the classroom, Weir (1993). If the sampling of the course book coverage is not satisfactory, the chance testees to demonstrate their knowledge of the language will be limited. Hughes (1989) notes that if a test lacks content validity, it results in a harmful wash back effect. It means, if a test is biased to

certain content areas in a syllabus at the expense of other content areas, those content areas which are less considered or not considered, will receive little attention or no attention from learners while practicing and studying. On the other hand, if all content areas in the domain treated in the classroom are given equal chance to appear in the test, there will be beneficial wash-back effect on teaching. That is learners will give balanced attention to all content areas of the syllabus during practice and study. Generally, the more representative samples a test contains, the more reliable will be the assessment of student's knowledge and ability. Based on this concept, this study is meant to meet the following objectives:

1.3. Objective of the Study

This study has one main objective. It attempts to discover whether or not classroom (teacher made) English language tests in Awassa College of Health Sciences are valid contentwise. It finds out the extent to which the coverage of sample tests is representative of the coverage of the course book.

1.4. Research Questions

1. Do tests in Awssa College of Health Sciences assess language Knowledge and skills defined by the objectives of the syllabus and course book units in an appropriate proportion?
2. If they assess, to what extent does their content coverage reflect the contents in the course book?

1.5. Significance of the Study

As a matter of fact, the result of any study must have importance to the users. The result of this study, therefore, will give the following significances:

- Helps English language teachers to get an insight about content validity of their tests.
- Creates awareness on School administrators and other concerned bodies if English language tests in the college are valid in terms of content

validity. As a result, school administrators can take appropriate measures depending on the result of the study.

- Serves as a resource material for future related studies.

1.6. Scope of the Study

There are three governmental Health science colleges in southern Nations Nationalities and peoples Region (SNNPR) which train health professionals in diploma (10 3) in different subjects (departments). However, this study was confined only to Awassa college of Health sciences. This is to make the study more manageable and because Awassa College of Heath Sciences uses sufficient number of employed English language teachers unlike other two which use guest teachers. As to tests, the study is limited to mid and end term tests of academic years 2006- 2008. Since the college has been established in 2004, there are no more tests to be considered. The three academic year tests are included in the study so as to get comprehensive information. The researcher wanted to study the content validity of tests given for third year students not for any special reason but randomly.

1.7. Limitation of the Study

In order to have more comprehensive information, it could have been good if the study had included sample tests from more colleges. But due to time and financial problems, this study was limited to one college.

CHAPTER TWO

2. REVIEW OF RELATED LITERATURE

2.1. What is a test?

A test has been defined by various scholars almost in a similar way except difference in ways of expression. To mention some of the definitions: Bachman (1990) defines a test as "... a test is a measurement instrument designed with explicit procedures to elicit a specific sample of individuals behavior." Similarly, Institute of curriculum Development and Research (ICDR) (1999) explains a test like this "A test is a systematic procedure of determining the extent to which instructional objectives are attained by pupil." Another Scholar, Good (1973) puts the definition of a test as "A test is a group of questions or tasks a student is to respond; the purpose being to provide a quantitative representation of pupils' trait it is designed to measure." Lastly, Brown (1994) defines it as a method of measuring individual's performance in a given area of a subject matter.

Thus, from the definitions given by the scholars mentioned above, a test can be conceived as a method or an instrument or systematic procedure which involves a group of questions (perhaps written or oral) that are mainly designed to elicit student's performance or to check the attainment of teaching objectives.

2.2. The link Between Teaching and Testing

Teaching language and testing language can not be viewed in isolation. They are the two components of the whole teaching learning process. In support of this idea, Heaton (1988) notes "It is virtually impossible to work in either field without constantly concerned the other." Likewise Dejene (1994) explains testing is like a vehicle to the well being of the total teaching- learning activity. Moreover, Venkatswaran (1995) states that students and teachers gain much

information from tests. The information gained in turn affects teaching positively. He says tests can help students to:

- develop positive attitudes towards their class
- Master the language
- Foster learning by their diagnostic nature

For teachers, tests offer lots of information. Good tests help teachers in that they answer an important question “Have I been effective in my teaching?” ‘ This question leads the teacher to diagnose his or her efforts made to teaching as well as help to answer questions below whose answers contribute much to the effectiveness of teaching- learning process

- Are my lessons on the right level?
- Am I aiming my instruction to low or to high?
- Am I teaching some skills effectively?
- What areas do we need to work on?
- Which areas need reviewing?
- Should I spend more or less time on the materials of mine? etc.

Madson (1983) also states that tests back up students in learning language. He puts his idea as follows:

A properly made English test can help create positive attitudes towards instruction by giving students a sense of accomplishment a feeling that the teachers evaluation of them matches to what he has taught them- - - tests also help students learn the language by requiring them to study hard, emphasizing course objectives and showing them where they need to improve.

Another scholar Guthrie (1984) in Nuru (1992) also shows the relationship between teaching learning, and testing as follows.

If the tests measure significant skills on highly valued subject matters that are included in the curriculum objectives of the schools, the potential for the genuine growth in children is high. However, if the tests are shallow, trivial or irrelevant to schooling and significant learning, curriculum and learning might a tail-spin (189).

In a more comprehensive manner, Institute for curriculum Development and Research (ICDR) (1999) describes benefits of testing to teaching as follows:

- Improve learning and instruction
- Help motivate learners to learn more as they discover their progress
- Identify learning difficulties that give learners opportunities to show the progress toward objectives;
- Help teachers to determine the effectiveness of their teaching aids, methods, techniques and learning materials;
- Provide educational administrators with adequate information about teachers effectiveness and the school program as a whole;
- Acquaint parents with their children's performance.

In general, it is likely to understand that tests of a right quality and standard can support teaching-learning. They make students develop a sense of achievement which by its turn motivates students for learning. In addition, they encourage students to plan their own strategies for better performance. For teachers, similarly, tests are mirrors to show weak and strong sides in the actual instruction. They provide a kind of information that helps teachers follow the right track of teaching. Underhill (1982) in Teshome(1995) explains

the strong bond between teaching and testing like this, As YE TEACH, So shall YE TEST’.

2.3. Types of Tests

Different purposes require different kinds of tests. Hence, depending on the purpose of testing, tests are classified into various types. They are categorized as proficiency tests, diagnostic tests, placement tests, progress tests and achievement tests.

According to Hughes (1989), “proficiency tests are designed to measure peoples’ ability or skills in a language regardless of any training they may have had in that language.” The content of such tests is based on a specification of what candidates have to be able to do in the language in order to be considered proficient for a particular purpose. In other words, proficiency tests help to provide a profile of an individual in terms of what he/she is or is not able to do. Moreover; they contribute to the prediction of individual’s academic success in future. They are not syllabus oriented. Tests conducted to employ translators, to admit graduate program students etc can be examples of proficiency test.

The second type of tests is diagnostic test. “Diagnostic tests are concerned with the persistent or recurring learning difficulties that are left not resolved by the corrective prescription of formative assessment,” (Thorndike 1977). It searches for the root causes of the learning problem. For example, what is the main cause that makes students confuse past perfect tense and simple past tense? Its main aim is to find out the real causes of learning problems and then to formulate a plan for remedial action.

The third is placement test. According to Harrison (1983) “placement test is a test designed to sort new students into teaching groups to make them start a course at approximately the same level as other students in the classes.” It is

not syllabus oriented test but is usually used to test general language ability of candidates rather than specific point of learning. According to ICDR, it focuses on questions:

- a) Does the learner possess the knowledge and skills to begin the planned instruction?
- b) To what extent is the learner ready to master the objectives of the planned instruction?

In general, placement test helps to get a clear picture of candidate's performance on a course before presentation of the instruction or the program.

Fourth, progress (formative) test. It is a continuous assessment that focuses on monitoring learning progress and diagnosing learning difficulties while teaching learning is under operation. It is usually conducted at the end of units, chapters or at any reasonable portion of a course. Like achievement test, progress test is purely syllabus based. It tests what is incorporated in a syllabus and treated in the classroom.

Progress test has a considerable contribution to teaching learning. To begin with, it helps testers (teachers) to identify their learner's weakness and strength at each segments of a course. Then they can take appropriate remedial actions on the areas where students are found weak at. It also assists teachers guide and make their work meaningful in that it provides ongoing feedback whether their teaching strategies are on the right track – make revise or keep on teaching strategies. Moreover, progress tests may inform teachers to adjust their pace of teaching. Heaton (1990) says. "Progress tests also act as a safeguard against hurrying on to complete a syllabus or a textbook regardless of what students are actually achieving or failing to achieve." Similarly, such a test is important for students because it makes them know how much they are progressing, avoid test anxiety since they are regular tests and make them hard worker (motivate them to study and practice day –to – day).

Fifth is an achievement test. Achievement test is generally referred to as syllabus- content test. Hughes (1989), notes that such test is firmly rooted to classroom experiences and syllabus contents. In the same way, Weir (1993) describes that since achievement test is viewed as an integral part of the learning process, it works best if it contains items and tasks about which students are familiar to. Such test takes place at the end of semester or school year for defined purposes.

The purpose of conducting achievement test is to measure student's progress in language skills in relation to the syllabus they have been following, Harmer (2001). Likewise, Weir (1993) says, "The over all aim of achievement test is to try to get students to use language receptively and productively that they have learned during the course." In short, achievement test is used to indicate how much successful the students have been in their language learning or how much the objective of the course is achieved. Student's scores from achievement tests can provide relevant information to teachers, students themselves, parents, curriculum developers and so on.

To attain the purposes underlined above, a good deal of attention should be given to sampling the items of the achievement test from contents specified in the syllabus or course book but which are realized in the classroom. Reasonable number of samples from each contents of the syllabus should be included in the achievement tests. Hughes (1989) states, "If achievement tests cover as much as the syllabus as possible, they tell how much students have progressed in the language." Problems are always likely to arise if the achievement tests do not contain satisfactory samples of a syllabus and if do not correspond to the teaching that preceded them. For example, weir(1993) notes, "If achievement test remains a test of grammar and vocabulary, teachers will spend much of their teaching time on these language areas at the expense of productive writing and speaking. Similarly, students will confine

their study to those areas which appear in classroom tests, not on anything else.” Teachers and students work hard on areas which will be tested.

Thus, it is very helpful if achievement tests incorporate adequate samples of language elements in the syllabus which have been treated in the classroom. Properly written achievement tests sampling correctly and proportionally from the course content can be indications to the learners which parts of the course need more attention and to course designers and teachers help to identify which part of the course have been effective and which have been ineffective. Then changes or modifications will be made on ineffective ones. Moreover, the more representative samples are included in the achievement test, the more reliable the assessment is likely to be and the more confidently we can generalize from performance on the tests.

2.4. Continuous Assessment

In formal educational settings, there are two main components of assessment systems. They are continuous assessment and end term assessment systems. Continuous assessment is defined by scholars from educational stand point in the following ways: “continuous assessment is a student evaluation system that operates at classroom level and is integrated with instructional process.” Capper (1996), in ICDR (1999). Similarly, yolo (1984) explains,” Continuous assessment is a method of evaluating the progress and achievement of students in educational institutions.” Hence, from these views we can understand that continuous assessment is ongoing type of assessment which helps facilitate learning as it often engages students in work by using various techniques of assessment.

Evidences show that today schools and universities are emphasizing the use of continuous assessment for it has various advantages in teaching- learning process. ICDR (1999) mentions the following advantages of continuous assessment for learners and instructional process.

For learners, it

- encourages participation in learning conditions.
- helps to analyze their level of knowledge and skills.
- determines their weakness and strength.
- motivates to regular practice and study.
- increases the reliability of their cumulative scores.

For instructional process, it:

- determines the extent to which the school is working towards achieving its set of objectives;
- provides information from which teachers can find insights in to their own effectiveness in teaching;
- finds out the degree to which the methods of instructional materials employed are up- to date and effective;
- gives incentives in learning process
- makes assessment of student's ability more reliable;

Generally, continuous assessment system brings teachers and students together in a cooperative endeavor to attain the objectives of classroom instruction. It checks student's progress and the effectiveness of teacher's instructional methodologies.

Scholars like Weir (1990) say for better construction of tests, it is strongly advisable to specify the objectives of the course in advance. However, all desired objectives of a course cannot be assessed with paper- and- pencil tests. For example, oral communication, psychometric skills, concept acquisitions, etc. Such content areas of a subject can be assessed by using continuous assessment system, because continues assessment system includes assessment techniques like observation. For example, students

speaking skills are best assessed via observation which involves well set checklists.

Concerning weighting, ICDR (1999) notes that schools can decide the weight for continuous assessment and end term (achievement) tests. It suggests 40-50 percent for continuous assessment. The indicated weight (40-50) can be gathered through different techniques of continuous assessment such as written test, observation, class work, home assignment, field work, portfolio, project work, laboratory work and so on.

In conclusion, continuous assessment like end term assessment, should give a proportional chance to the sampling of objectives or contents of a given course. ICDR (1999) emphasizes “continuous assessment procedure should include a representative sample of instructional objectives treated in the class.” So; both continuous assessment and end term assessment tests are assumed to test contents that are presented in a syllabus as long as they are treated in the classroom.

2.5. Qualities of a Good Test

“Tests can affect quality and quantity of learning” Kohonon (1999), in Arnold (1999). The effect of testing on learning can be beneficial if a test involves the three major criteria of quality test. These criteria are practicality, reliability and validity.

2.5.1. Practicality

Test practicality refers to the consideration of financial capacity, administrative capacity, and scoring capacity of the target setting, Harrison (1983). A test should be economical in terms of money and time, easy to administer and score. Since practicality, reliability and validity are dependent on each other, tests can be valid and reliable if they are practicable, Weir (1990). Thus, a test is said to have a required quality if it is practical.

2.5.2. Reliability

Reliability is another vital criterion for the goodness of a test. It refers to the extent to which test scores are consistent when a group of students take it in two occasions at about the same time (Alderson, Calhan and Wall 1995). A test is said to be reliable when test scores form the same types of tests (test- retest or parallel test) given to the same group of students without referring to any notes and get support is roughly similar. If the scores are consistent then the tests are regarded as accurate values of student's achievement. Hence, statements about students' performance can be acceptable. Bachman (1990) notes that a test score even should be consistent when a test is conducted to the same group after some period of time gap, in different test forms and marked by different markers. If there is variation between the scores of the tests in the two occasions, the test is seen as unreliable test (Ibid). So, the scores may not be dependable to make decision about student's performance level.

Reliability has a strong relationship with validity. They are interdependent qualities of any test. Weir (1993) states, "Validity is the first point in test design. But for the test to be valid, it must also be reliable." Similarly, Underhill (1991) says, "I prefer to see reliability as a specific type of general validity. A test can be valid unless it is reliable." Moreover, Alderson, Clahpan, and wall (1995) explain that unreliable tests cannot be valid but valid test can be unreliable due to factors external to the test itself. Such factors are testing conditions, scorers etc. From these scholars point of view, its likely to conclude that a test to be good, it should involve validity and reliability simultaneously.

There are various factors that can affect the reliability of a test. These factors lie in test itself, test raters, testees and test area conditions. Each factor is illustrated as follows.

- Test itself: includes length of the test (too long or too short), instruction clarity, adequate sampling of test items and so on.
- Testees: includes illness on the exam day, motivation towards a test, sleepless night etc.
- Test rater: includes subjective techniques of marking etc.
- Test area conditions: includes light, noise, temperature, test administrators etc.

Thus, to establish the reliability of a test, testers should strive to minimize the above mentioned factors. If so, test scores can be regarded as real reflections of students' language proficiency level.

2.5.3. Validity

Test validity concerns with the degree to which a given test or any of its component parts should measure what it is intended to measure, not anything else, Weir (1993). Validity is usually taken as a crucial frame of reference for a test writer to come up with a quality test (Ibid). Every item in a test should test in a clear way what a tester has intended to test if he/she is to say anything meaningful about testees' ability in using a language. Such a test is said to be a valid test.

Validity is divided into four main kinds. These are face validity, concurrent validity, predictive validity and content validity. Each of these validity aspects are further explained as follows.

2.5.3.1. Face Validity

Face validity is concerned with a test's surface level credibility or public acceptability, Ingram (1977), in Alderson, Claphan and wall (1995). Similarly, Harrison (1983) describes "Face validity is concerned with what teachers and students think of the test." Moreover, Underhill (1991) notes "Face validity is an aspect of validity achieved from the questions like on the face of it, does it

look like a reasonable test? Do the people who use the test think it is a good test?”

As to researching face validity, Underhill (1991) suggests the following way:

Clearly, the best way of researching face validity is to question different people who come to contact with the test. Learners generally produce very informative and objective comments about tests, irrespective of their own personal performance; any doubts about this can be easily checked by comparing the general tone of their comments with their scores.

In conclusion, face validity is not empirical validity. It simply deals with the appeal of the test to candidates, teachers, family, members of the public etc. It doesn't need experts to judge for it deals with surface levels of a test such as number of items, formats, font size, space between lines etc.

2.5.3. 2. Concurrent Validity

Concurrent validity is another branch of validity. According to Alderson, Claphan and Wall (1995), it refers to a test's validity which is obtained by making comparison between scores of one test with the scores of some other test on the same subject given to the same group of testees. Here, the other test refers to the parallel version of the same test or a test in different form but the same content. As Teshome (1995) notes concurrent validity is indicative to the validity of the new test. If the two tests have a high correlation, the new test is valid provided that the first test is highly dependable assessment of testees ability.

Underhill (1991) expresses that the two tests are said to have perfect concurrent validity if the correlation coefficient is 0.9 or more. This suggests that the tests are measuring the same thing. However, the two tests are said to

have low concurrent validity if the correlation coefficient is 0.4 or less- suggests that the two tests are measuring different things. The vast majority of correlation coefficients fall between these figures, and no single conclusion can be drawn from them. Actually, the value of correlation coefficient falls between +1 and -1. Logically thinking, the concurrent validity coefficients are high if the two tests are tests of the same skill, Underhill (1991). For example, two oral test scores correlate more highly than one oral test score with multiple choice grammar test score.

2.5.3.3. Predictive Validity

Such validity is a kind of validity which is obtained by comparing a test scores with some other measures given to the same candidates some time after the first test, Alderson, Claphan and wall (1995). Predictive validity helps to know that how predictive is the first criterion test to the achievement in the later learning. For example, how predictive are Ethiopian General School Leaving Certificate Examinations (EGSLCE) Scores to college grades.

Predictive validity and concurrent validity are similar in that both of them involve mathematical formulae for the computation of validity coefficients. But the time gap between the first test and the second test in the case of predictive validity is wider than the time gap in the case of concurrent validity. Correlations in predictive validity are harder to interpret with confidence than concurrent validity correlations because of all the influences that can affect the learners' abilities and performance between the two test measures due to the longer time interval between the two tests.

2.5.3.4. Content Validity

Many testing scholars have given their own definition to content validity. All definitions are similar in their central focus except the way they are worded. The following are some of the definitions by different scholars.

Harrison (1983) and Alderson, Claphan, and wall (1995) define content validity as a test's ability to include or represent all of the contents of a particular domain (syllabus) in a proportional manner. Similarly, Bachman (1990) defines "Content validity is the extent to which the tasks required in the test adequately represent the behavioral domain in question." Moreover, Weir (1990) describes content validity as "The extent that tests sample as widely as possible relevant, critical and communicative items from the syllabus in order to have a beneficial wash- back effect on teaching."

As we see, all definitions forwarded above to content validity are almost similar in that they express content validity as a test's extent of coverage of contents available in the syllabus by drawing representative samples in a proportional manner. According to Anastasi (1982) in Weir (1990), whether a test covers representative samples of language contents from a given syllabus can be determined by the systematic examination of the test's content.

According to Alderson, Claphan, and Wall (1995) and Underhill (1991), "Common way of content validation is to analyze the contents of the test and compare with the statements of what the content ought to be. The statement may be syllabus or domain specification." Hence, to analyses content validity of tests, the coursebook and sample tests' contents should be quantified in terms of frequencies of practice items in the coursebook and test items in sample tests. Oller (1979) in Alemu (1983) notes that analysis of language contents can be quantified by virtue of scoring procedure that counts certain language units in the surface form of the discourse. To ease the analysis of the test content, using test content specification table is advisable, Hughes (1989). To ease the analysis of the test content, using test content specification table is advisable, Hughes (1989).

2.5.3. 4.1. *The Importance of Content validity to Language Learning*

Though all aspects of validity have significance in teaching learning, content validity is the most relevant one because it is a means to check the attainment of objectives of each contents of the syllabus or the universe of a certain domain, Hughes (1989). If a certain test is valid content coverage wise, it provides relevant information to the concerned bodies that how much students have progressed in each contents or the syllabus. In other words, It tells the tester or/and other beneficiaries of a given test to what extent the objectives of the course have been moving towards the desired goal.

Besides, tests which involve a proper content validity trigger students to study hard each contents of the course. Otherwise, students will avoid studying and practicing those language areas which do not appear or appear less in tests. They give a considerable attention to those language areas which they are going to be tested. In favor of this, Kohonm (1999) notes,

It is known that anticipation of testing procedures has a wash back effect on learning; learners prepare for examinations and organize knowledge in memory in the light of how they are going to be tested. ... evaluation thus affects both quality and quantity of learning. Therefore, it needs to be examined in terms of both the learning process and the outcomes of the learning.

In a similar way, Weir (1993) says, “students could perform well in language areas or skills where tests are emphasized; that is to say, students study and practice more language areas to which more emphasis is given during testing.” By this, Weir implies that considering all contents of syllabus proportionally during testing, facilitates language learning. Lastly, scores gained from tests which involve satisfactory sampling (content validity), can help to approximate students’ actual performance level. That is they guarantee to draw acceptable

statements about learners' proficiency status in a language in a particular grade level.

To sum up this section, if tests contain satisfactory samples from each contents areas of a given syllabus, they can have a positive influence on teaching learning. It is important, therefore, that tests should sample as widely as possible test items from each contents of the domain. As a result, decisions that one makes about students' performance level and the program or syllabus, will be more acceptable. Decisions, based on scores gained from tests which are poor at content validity, are likely to be imprudent (less acceptable). Tests with appropriate content validity are helpful to increase quality and quantity of learning, provides valid and reliable information about students' performance and shape the teaching procedure.

2.5.3.4.2. Guidelines to Establish Content Validity

As it was discussed earlier, content validity is a vital feature of any test. Multitude of scholars share the idea that tests of a good content validity facilitates learning a language and other subjects. Hence, test writers should give a considerable care while writing tests. Anastasi (1982) in Weir (1990) suggests some very important guidelines that help test constructors to establish content validity of tests. These guidelines read as:-

- the behavior domain to be tested must be systematically analyzed to make certain that all major aspects are covered by the test items and in the correct proportion.
- the domain under consideration should be fully described in advance, rather than being defined after the test has been prepared.
- content validity depends on the relevance of individual's test response to the behavior area under consideration, rather than on the apparent relevance of item content.

Thus, language test writers are advised to undergo through these guidelines in advance in order to produce tests which involve adequate content validity since content validity is one of the key promoters of learning language. If the test designers do not follow the guidelines suggested, it is very likely that they forget to include relevant language elements into a test. Even the included ones will lack proportionality. Some language areas will be dominant in the test.

CHAPTER THREE

3. METHODOLOGY

3.1. Data Sources

The primary aim of this study was to assess the content validity of English language tests in Awassa College of Health Sciences- with reference to English language tests Conducted to third year students in three consecutive years – 2006, 2007, and 2008. For the study, the main data sources were the course book designed to third year students and sample test papers prepared for the same group of students in the said years.. Moreover, English language teachers in the college were used as sources of some information that are relevant to the validity of the study. They were asked questions concerning whether they cover all course book units, use continuous assessment form. This helps especially to know if they test listening, speaking and writing.

3.2. Sampling Materials

As was mentioned earlier, materials that this study has utilized as data source are course book and sample test papers. The course book was chosen randomly, whereas, the sample tests were selected by using comprehensive sampling technique (a technique that considers all population as study subjects). The college was established in 2004. Therefore, it was only three times (in 2006, 2007 and 2008) English language final classroom tests were provided for third year students. So, all available final achievement classroom tests (one in each year), were included into the study. Similarly, one progress test from each year were considered in the study because the modes of evaluation in the college is through continuous assessment, which includes one progress test (mid test) and final achievement test. There is not any more progress test to be sampled. Instead of including one year or two years of tests in the study, the researcher to considered progress and final achievement tests

in three years because it guarantees to arrive at a better conclusion about the content validity of tests in the college.

Similarly, teachers for unstructured interview were selected by using comprehensive sampling method. There were four English language teachers in the college. All of them were considered to the interview. They have their first degree and served two to four years in the college. All of them involved in writing both progress and achievement classroom tests in 2007 and 2008. But two teachers only participated in preparing the tests in 2006. This shows that information that teachers gave have direct attachment with the sample tests of the said years.

3.3. Data Collecting Procedures

In order to gather data for the study, the following procedures were followed:

3.3.1. Course book Analysis

First, the contents of the course book were classified into major headings such as reading comprehension, vocabulary, listening, speaking, writing, grammar, punctuation and pronunciation. Then the total number (frequency) of items in practice exercises in each major course book content area were determined and put in simple figures and percentages in a table. To increase the reliability of classification of content areas of the course book and determining frequencies of items of practice in to each content areas of the course book, two English language teachers in the study college were made to participate. Moreover, different language people, for example, graduate program students in the department of TEFL were consulted.

3.3.2. Interview to Teachers

Next to course book content analysis, unstructured interview (informal conversation) was conducted to English language teachers in the college- in order to know whether they:

- a) cover the whole units of the course book in each year. This is to ensure if the contents of the sample tests were sampled from the whole course book coverage or not. It has significant implication on the result of the study.
- b) use continuous assessment form (method). If they use, which language skills or sub-skills they test through it? How many times? Out of what percent? This helps to know if speaking, listening, writing and other skills are assessed through continuous assessment in addition to the sampled progress and achievement tests. Especially, to know if speaking and listening skills which are not usually expected in paper- and –pencil tests were tested or not.
- c) have taught English for third year students in 2006, 2007 and 2008. This helps to assure if teachers responses under **a** and **b** work for tests in 2006, 2007 and 2008.

Unstructured interview was preferred because the researcher feels that teachers will give reliable information during informal interview. In formal interview, they may become more careful and be reserved to give some important information, especially what they do not do.

3.3.3. Sample Test Paper Analysis

After course book analysis and administration unstructured interview to teachers, progress and achievement sample test papers of the years 2006, 2007, and 2008 were collected. The collected sample test papers were analyzed in the same way as course book analysis. Headings such as reading comprehension, vocabulary, grammar, listening, speaking writing, punctuation and pronunciation were determined as major content areas of the sample test papers. Next to this, test items in the sample tests of each year, were classified under major content areas which are already determined and then summarized in a table in frequencies. As was done in the case of the course book, to maximize the reliability of the classification of test items into

appropriate content areas, the same English language teachers were made to participate. In addition, graduate program students in the department of TEFL were consulted.

In the end, the degree of relationship between the contents of the coursebook and sample test papers was computed by using Sidney Siegel's formula, $C = \frac{\sqrt{x^2}}{\sqrt{N+x^2}}$

Where, C= the value of relationship

N = grand total (the sum of row totals across each contents)

X^2 = (chi-square) the sum of observed values (frequencies) minus expected values (frequencies) the whole square and then divided by the expected values across all categories (content areas in the course book and sample test papers).

CHAPTER FOUR

4 RESULTS AND DISCUSSION

4.1. A Brief Overview of the Course Book

Presently, the coursebook, which this study focuses on, is under implementation in all colleges that train learners in a diploma level and in preparatory schools for grade 12 students. It was first produced in 1999 and revised in 2000. The course book has seven units each of which is classified into seven sections. The sections in each units are reading comprehension, vocabulary, note-making and summary, speaking, listening, writing and grammar.

The course book is accompanied by a syllabus. In the syllabus, general objectives of the course as well as periods allotted to each content areas are indicated. As Weir (1993) notes testing should correspond with the objectives of the course, methods of teaching, and exercise formats. This is to say, testing and objectives of any course should go hand in hand, hence, it will be important to indicate the objectives of the course here.

- General objectives: The main objectives of the course are to make students.
 - Use the English language spontaneously and flexibly for oral communication in an academic environment and in other situations where the use of English is appropriate.
 - Understand different types of written materials in English that they will study at school.
 - read efficiently, critically and make discriminating judgments
 - Produce the kinds of writing which will be expected of them in their chosen subjects, e.g essay, reports...

- Employ the study skills which they will need in order to operate effectively as tertiary level students.

Thus, tests are expected to test these general objectives in one way or another.

4.2. Analysis of Content Areas of the Course book

English Course book for third year students were analyzed in terms of major content areas and frequencies of practice items (questions for practice) in each major content area. Major content areas were determined as reading comprehension, vocabulary, grammar, listening, punctuation, speaking, writing and pronunciation. The number of times that the items of practice (questions for practice) in each content area of the coursebook, was put in frequencies. For example, punctuation practice exercises in the course book have 30 items (questions for practice). So, it means that punctuation as a content area has a frequency of 30 in the course book.

In detail, the following table (table 1) demonstrates major content areas of the coursebook with their corresponding items of practice in frequencies.

Table 1: content area of the course book and frequency of items of practice

No	Content area	Frequency of items	%	Remark
1	Reading comprehension	107	19.5	
2	Vocabulary	165	30.1	
3	Grammar	205	37.4	
4	Listening	31	5.7	
5	Punctuation	30	5.5	
6	Pronunciation	10	1.8	
	Total	548	100	

As we can see from table 1, the content areas of the coursebook are classified into eight possible sections. Moreover, the frequency of items in practice

exercises under reading comprehension, vocabulary, grammar, listening, punctuation and pronunciation are indicated in the table. But for writing and speaking skills frequencies were not determined because it is difficult to put items or activities for them quantitatively (in frequencies) for they are very general. For instance, speaking tasks are debating, describing and so on. Similarly, writing activities are like writing summary, writing composition etc. Hence, they are difficult to express in frequencies for comparison purpose. Spelling exercises are included under vocabulary. But punctuation exercise (item) is treated separately though it is an aspect of writing.

Table1 shows that from the total of 548 practice items in the course book, reading comprehension contains 107(19.5%) items, whereas, vocabulary and grammar make up 165(30.1%) and 205(37.4%) respectively. Listening, punctuation and pronunciation constitute 31(5.7%); 30(5.5 %) and 10(1.8%) respectively. As we can see the table, grammar and vocabulary received more emphasis. Next to grammar and vocabulary, reading comprehension is emphasized with 107(19.1%) items. The rest areas seem to have less emphasis. There is significant difference in the number of items of practice among the content areas. For example, between grammar and reading comprehension, there is a difference of 98 items, between vocabulary and listening 134.

As was said earlier, activities, on speaking and writing are very difficult to express quantitatively because they are very general. However, it is fair to see them in terms of period allotted to them, because number of periods given to contents can be an indication of emphasis given to that particular content area. Based on a syllabus, table 2 shows a summary of periods given to each major content areas of the course book.

Table 2: Summary of contents of the syllabus on the basis of period allocated

No	Content area	Total time	%	Remark
1	Reading comprehension	28	14	
2	Vocabulary	27	13.5	
3	Grammar	29	14.5	
4	Listening	28	14	
5	Speaking	44	22	
6	Writing	44	22	
	Total	200	100	

Table 2 indicates that from the total of 200 periods, speaking and writing have taken the highest share. Each of them covers 44(22%). Reading comprehension and listening have been given equal periods. That is 28(14%) each. Vocabulary and grammar have 27(13.5) and 29(14.5%) shares respectively. From the point of view of the periods distributed, we can say that speaking and writing are given more emphasis than others. They, together constitute 88(44%) of the total periods. The rest four have been given 112(56%) periods.

4.3. Analysis of sample test papers

As Underhill (1991), Alderson, Claphan and Wall (1995) note that to determine whether a certain test is valid contentwise, the first step is to categorize the syllabus (course book) in to major content areas. The second step is to determine the number (frequency) of practice items in each content area of the course book and the frequency of test items in each content areas of the test paper. Hence, sample test papers on which this study focuses are classified into content areas in six tables. Within the same tables, a class of each test items is indicated (see appendix 1 to 6).

The horizontal rows of each table shows major content areas of sample test papers, whereas, the vertical rows contain question numbers of each items in the test papers. The class of each item in the sample test papers is indicated by the thick mark ($\sqrt{}$) within the tables. For example, in appendix 2, question numbers 7, 13,14,21,23 and 24 are marked using a $\sqrt{}$ by the vocabulary row because they are testing meaning of words.

The analysis of the distribution of different content areas of the sample test papers is summarized in figures and percentages in the following table

Table 3: – Summary of total frequency of items in various content areas in sample test papers by content areas and test year

Content area	Frequencies by test year							
	Total items in 2006	%	Total items in 2007	%	Total items in 2008	%	Total	%
Reading comprehension	9	26.5	5	11.4	9	18.8	23	18.3
Vocabulary	7	20.6	17	38.6	25	52.1	49	38.9
Grammar	14	41.2	16	34.4	9	18.8	39	31
Pronunciation	0	0	0	0	0	0	0	0
Punctuation	4	11.8	6	13.6	5	10.4	15	12
Total	34	100	44	100	48	100	126	100

As it is demonstrated in table 3, in the academic year 2006, from the total of 34 items, grammar covered 14 (41.2%) which is 50% of the total percent given. Reading comprehension comprised 9(26.5%) items. vocabulary and punctuation constituted 7(20.6%) and 4(11.8%) items respectively. However, pronunciation was totally ignored. There is not any item that represents it.

As the distribution shows, grammar was dominant with 14(41.2%) as it is dominant in the course book with 205 (39.7%) items of practice. The next dominant content area in this test distribution is reading comprehension but it is third in rank in the course book. Vocabulary became third but it is second in the course book. So, vocabulary and reading comprehension are reciprocated. Punctuation is well represented as compared to reading comprehension and vocabulary. In the course book, it has coverage 5.8% but in this tests 11.8% coverage.

Concerning tests in 2007, from the total of 44 items, vocabulary and grammar covered 17(36.6%) and 16(36.4%) items respectively. Reading comprehension constituted 5 (11.4%). punctuation 6(13.6%). But pronunciation was totally rejected. Vocabulary has taken the highest coverage with 17(38.6%). Next to it is grammar with 16 (36.4%). They were reciprocated unlike their coverage in the course book. Similarly, punctuation and reading comprehension are reversed. Punctuation preceded reading comprehension. The former is 6 (13.6%) but the latter is 5 (11.4%).

In 2008, from the total of 48 items, vocabulary covered 25(52.1%). Grammar and reading comprehension were given equal emphasis. Each of them covered 9(18.8%). Punctuation constituted 5(10.4%). However, pronunciation was totally ignored.

In the distribution, vocabulary has taken the highest share with 25 (52.1%) items. It has covered more than half of the total distribution. Unlike their coverage in the course book, grammar and reading compression have equal rank. In the course book grammar is twice reading comprehension, with 205 and 107 items respectively. Grammar was insignificantly represented in the sample tests, however.

So far the two sets of important data have been obtained by analyzing contents of the coursebook and sample test papers in terms of frequencies of items of practice and test, the extent of relationship between the two observations can be computed. Data from coursebook analysis are indicated in table 1. Similarly, data from sample test paper analysis are presented in table 3. The total frequencies of items for both observations are put in figures and frequencies.

In order to determine match/mismatch between the contents of coursebook and test papers, Siegel's formula was used. The formula is

$$C = \sqrt{\frac{x^2}{N + x^2}}$$

Where, C= the value of relationship

N= grand total (the sum of all row totals)

$X^2 = \sum (K_i)$ is the summation of observed value minus expected value the whole square divided by expected value.

Where, observed value refers to frequencies occurred in each category (language content areas), whereas, expected value refers to frequencies that were expected to occur in each category.

To find x^2 , three steps are followed. These are:

Step 1: Determine the content areas of the course book and sample test papers. Then put the items in each content area in frequencies. Then after, determine row and column totals. The row total is the sum of item frequency in the course book and in test papers across each content area. But the column total for the coursebook is the sum of frequencies of practice items in all content areas of the course book (see table 4 below), however, the column total for the sample test paper is the sum of frequencies of test items in each content area of test papers. See table 4 below.

Step 2: compute the expected values for each observed values using the formula, expected value = $\frac{\text{row x column total}}{\text{Grand total}}$.

Grand total

- Step3: (a) Subtract expected values from observed values
 (b) Square the differences
 (c) Divide the squared differences by expected value
 (d) Add all quotients.

Therefore, the sum of all quotients gives χ^2 . The relationship value or degree of relation C, is computed by using the formula

$$C = \sqrt{\frac{\chi^2}{N + \chi^2}}$$

Table 4: Total frequencies of content areas of course book and sample test papers

Content areas	Frequencies in course book	Frequencies in test papers	Total frequencies
Reading comprehension	107	23	130
Vocabulary	165	49	214
Grammar	205	39	244
Pronunciation	10	0	10
Punctuation	30	15	45
Total	517	126	643

Table 4 shows that the column total for the coursebook and sample test papers is 517 and 126 respectively, whereas, the grand total is 643. Moreover, table 4 demonstrates row totals. For instance, the row total for reading comprehension is 130.

Table 5– The degree of relationship between the course book and sample test papers

	Observed	Expected value	X²
Frequencies in course book	107	$\frac{130 \times 517}{643} = 104.5257$	$\frac{(107-104.5257)^2}{104.5257} = 0.0586$
	165	$\frac{214 \times 517}{643} = 172.0653$	$\frac{(165-172.0653)^2}{172.0653} = 0.2901$
	205	$\frac{244 \times 517}{643} = 196.1866$	$\frac{(205-196.1866)^2}{196.1866} = 0.3959$
	10	$\frac{10 \times 517}{643} = 8.0404$	$\frac{(10-8.0404)^2}{8.0404} = 0.4776$
	30	$\frac{45 \times 517}{643} = 36.1820$	$\frac{(30-36.1820)^2}{36.1820} = 1.0562$
Frequencies in sample test	Observed	Expected	X²
	23	$\frac{130 \times 126}{643} = 25.4743$	$\frac{(23-25.4743)^2}{25.4743} = 0.2403$
	49	$\frac{214 \times 126}{643} = 41.9347$	$\frac{(49-41.9347)^2}{41.9347} = 1.1904$
	39	$\frac{244 \times 126}{643} = 47.8133$	$\frac{(39-47.8133)^2}{47.8133} = 1.6245$
	0	$\frac{10 \times 126}{643} = 1.9596$	$\frac{(0-1.9596)^2}{1.9596} = 1.9596$
	15	$\frac{45 \times 126}{643} = 8.0180$	$\frac{(15-8.0180)^2}{8.0180} = 4.3399$
Total			=11.6331

$$C = \sqrt{\frac{x^2}{N+x^2}}$$

$$C = \sqrt{\frac{11.6331}{643+11.6331}}$$

$$C = \sqrt{0.01777}$$

$$C = 0.13$$

As a table 5 shows, X^2 was determined. It has been found 11.6331. as result the extent of relationship between course book and sample test papers has been computed. It is 0.13.

4.4. DISCUSSION

By using Siegel's statistical formula, the extent of relationship between the contents of the course book and sample test papers has been determined. In other words, the question how much do tests in Health science colleges reflect the course book coverage is answered. The result is that the correlation coefficient value C is 0.13.

According to Siegel and Spearman, two observations are said to have a perfect relationship if coefficient value is 1 (one). But if correlation value reads less than 0.5, there is weak correlation between the two observations. The two observations in this study are contents of the course book and sample test papers. Therefore, the result of this study, the value 0.13, tells us that there is hardly any relationship between the two observations (contents of the course book and sample test papers). This means that the coverage of sample tests does not adequately reflect the course book coverage. In other words, sampling the course book coverage has been made improperly.

As we see in table 3, grammar has a big share in the course book. However, in sample test papers, in 2007 and 2008, it was given second position. Conversely, vocabulary which is second dominant in the course book was given first rank in sample test papers. So, grammar and vocabulary were reciprocated. Moreover, in 2007, punctuation outnumbered reading comprehension unlike in the course book. That is, punctuation was given a high coverage in the sample test papers than reading comprehension which is unfair allocation. When we consider the coverage of grammar and reading comprehension in 2008, they were given equal coverage which is again unfair. Both constituted 9(18.8%). However, in the course book their coverage is significantly different. Grammar is nearly twice the coverage of reading comprehension. The other problem is that pronunciation was ignored in all sample tests but it has a coverage of 10(1.9%) in the course book. This has affected validity of the tests, especially, content validity.

Hence, it is likely to say that there is disparity in allocation of percentages (samples) to the content areas in the coursebook. Such disparity or proportional representation of content areas of the course book into tests negatively affects the validity of tests. If important areas in the course book are not reflected in the classroom progress as well as achievement tests, the tests remain weak in different aspects. For example, inadequacy in sampling from the course book coverage affects teaching learning in that the test can't check whether students have got progressed in certain content areas or objectives, limits students opportunities to demonstrate their knowledge, demotivate students from further learning, makes students results unreliable and so on.

Therefore, tests should reflect different segments of the course book proportionally in order to meet their purposes such as assuring students progress in each objectives or content areas of syllabus, motivating students

for learning, giving opportunity to students identify their weak and strong sides, providing teachers with feed backs on their methodological issues etc.

With reference to testing speaking, writing and listening skills, the study has come up with the following points. Writing skills seem to have received relatively a better consideration than speaking and listening skills. As is observed in the three year sample test papers, there are about twentyone items on writing skill. Eleven of them were guided writing tasks and one was free writing task. However, the other nine are simple concepts about writing. They were limited response questions in multiple choice forms. They do not involve pen and paper work. In continuous assessment, 25% value was given to writing skill. Teachers reported that they usually give three to four tasks on writing skills through class work and home assignment. They encourage students to write paragraph and essay level texts on selected texts. In general, relatively speaking, testing writing skill was given a better attention compared to attention given to testing speaking and listening skills.

When we come to testing speaking and listening skills less emphasis has been given to them. As teachers say, they assess speaking via continuous assessment out of 20%. This percent value was collected from two presentation tasks. But it is unlikely to think that twice presentation can test students' speaking skill. It is too limited. Learners should be given repeated exposure to tests of speaking because recurring exposure can approximate the actual students' performance on oral skills. The time coverage that the syllabus allocated to speaking is high. It covers 44(22%) periods. This period is nearly twice of the periods given to other skills except for writing which is 44(22%). This shows that the syllabus has given much emphasis to speaking skill. However, the assessment is very limited. Only two opportunities were given to learners to assess their ability to speak.

In the course book listening skill has 31(5.7%) items of practice and 28(14%) periods. Despite the coverage indicated, it was totally ignored in testing. This

affects student's attention to listening activities in the classroom. Students are usually eager to learn what appears in the test as Weir (1993) emphasizes. As to teaching, teachers reported that they teach listening by dictating texts presented in teacher's book. But against principles in testing such as 'as you teach you shall test', testing listening was ignored.

CHAPTER FIVE

5. SUMMARY, CONCLUSION AND RECOMMENDATIONS

5.1. Summary

Comparison between contents of the course book and sample test papers was made after due analysis of each content. This was to find out whether the coverage of the sample test papers adequately represents the coverage of the course book. The finding of the study shows that the relationship (in terms of content coverage) between the contents of the two observations is weak in that the allocation of the contents of the course book in sample tests was not done in proportional manner. Some content areas which received high emphasis in the course book were given less attention in the sample tests. For example, grammar was given more emphasis in the course book than emphasis given to vocabulary. However, in the test papers, vocabulary received more emphasis. They were reciprocated.

Moreover, sample allocation among reading comprehension, grammar and punctuation showed inconstancy. For example, in 2007, punctuation was given more emphasis than reading comprehension unlike their emphasis in the course book. Reading comprehension covers more than threefold of punctuation coverage. Similarly, grammar and reading comprehension were given equal coverage in sample test papers in 2008, which is unfair. On top of these, some content areas which have significant coverage in the course book were not totally tested. Listening skill and pronunciation were absent in tests.

Thus, due to disparities discussed above, the extent of relationship between the contents of the course book and sample test papers has been found to be weak. It is 0.13.

5.2. Conclusion

This study has made an assessment on the content validity of English language tests in Awassa College of Health Sciences (ACHS). In other words, the study was attempted to determine the relationship between contents of the course book and sample test papers. Based on the data obtained from course book and sample test paper analysis as well as teachers' responses, the study shows that the coverage of the sample tests doesn't adequately represent the course book coverage. The language skills and knowledge included in the course book were inadequately incorporated into the contents of the sample test papers. That is, some language content areas treated in the course book were not represented in tests proportionally. Percentage allocation was not fair and some important contents were not totally tested despite their significant coverage in the course book. This shows that the sample tests remained weak at assessing the learner's knowledge of language proficiency in each area of the syllabus. Learners' opportunities to show their progress in every segments of the syllabus were restricted as the test contents were unrepresentative.

Due to these deficiencies, it is concluded that English language tests in Awassa College of Health Sciences are found to be weak at content validity, because they did not reflect all contents in the syllabus proportionally, as well as, they avoided important language skills and knowledge from the content of the test. As a result, the extent of relationship between the contents of the course book and sample test papers has been found to be 0.13 which means there is hardly any relationship between the two observations.

5.3. Recommendation

The result of this study shows that the content validity of sample tests in Awassa College of Health Sciences is very low. There is hardly any

relationship between the coverage of the course book and sample test papers. Hence, the writer of this thesis would like to forward the following recommendations which help maximize the content validity of the English language tests in the college.

- Before writing tests, teachers should make course book content analysis on the basis of coverage of each content areas of the course book. This helps them to determine the share of each content area in test papers.
- The basic science unit of the college should set committee that approves tests before administration.
- The concerned authorities should facilitate in-service training or workshops to teachers on language testing.
- It would be helpful if universities and colleges offer a course on language testing in undergraduate level.

BIBLIOGRAPHY

- Alderson, J., Claphan, C. And Wall, D. (1995). **Language Test Construction and Evaluation**. Cambridge: Cambridge University Press.
- Alemu Tsegaw (1983). "An Assessment of Grade Six and Eight English National Examinations". (Unpublished M.A. Thesis) Addis Ababa University.
- Atkins, J., Hailom, B, and Nuru,M. (1996). **Skills Development Methodology**. (Part 2). Addis Ababa: Addis Ababa University Press.
- Bachman, L. (1990). **Fundamental Consideration in Language Testing**. London: Oxford University Press.
- Berhanu Moges (2004). "Teachers Assessment of students Performance: Continuous Assessment" (Unpublished M.S. Thesis) Addis Ababa University.
- Brown, H (1994). **Principles of Language Learning And Teaching**. San Francisco State University: Prentice Hall Regents.

- Dejene Leta (1994). **Testing and Its Practical Application: Testing Reading In Focus (Unpublished Seminar Paper)**. AAU: Department of Foreign Language and Literature.
- Fallik, F. and Brown, B. (1983). *Statistics for Behavioral Sciences*. The Dorsey Press: Homewood.
- Good, C. (1973). **Dictionary of Education**. New York: Mc. Gram Hill.
- Harmer, J. (2001). **The Practice of Language Teaching (3rd ed)**. Longman: Parsons Education Limited.
- Harrison, D. (1983). **Language Testing Handbook**. New York: Cambridge University Press.
- Heaton, J.B. (1988). **Writing English Language Test**. New York: Longman Group Limited.
- Heaton, J.B. (1990). **Classroom Testing**. New York: London.
- Hughes, A. (1989). **Testing For Language Teachers**. Cambridge. Cambridge University Press.
- ICDR (1999). **Teacher Handbook**. Addis Ababa: Ministry of education.
- Kifle (Kebede (1995). "An Assessment of Content Validity of High School English Language Tests." (Unpublished M.A. Thesis) Addis Ababa University.
- Madson, H. (1983). **Techniques in Testing**. Hong Kong: Oxford University Press.
- Nuru Mohammed (1992). "Level of Questions: A Description of Textbook and Examination Questions in Higher Secondary Schools" (Unpublished M.A. Thesis). Addis Ababa University.
- Teshome Demisse (1995). "The Construction and Validation of Tests in English for Tertiary Education." (Unpublished Ph.D. Dissertation). Addis Ababa University.
- Thorndike, R.L and Hagen, E.P (1977). **Measurement and Evaluation in Psychology and Education**. New York: John Wiley and Sons.

- Tibebe Alemayehu (1992). "The Predictive Validity of Ethiopian School Leaving Certificate Examination English and the Integrative Tests: Comparative Study." (EJE, Vol. XIII, No 1).
- Tylor, R. (2000). **English for Ethiopia**. Addis Ababa: Ministry of Education.
- Tylor, R. (2000). English Syllabus for Grade 12. Addis Ababa: Ministry of Education.
- Underhill, N. (1991). **Testing Spoken Language**. Cambridge: Cambridge university press.
- Venkateswaran, S. (1995). **Principles of Teaching English**. Delhi:Vikas Publishing Pvt. Ltd.
- Weir, C. (1990). **Communicative Language Teaching**. New York: Prentice Hall.
- _____ (1993). **Understanding and Developing Language Tests**. New York: prentice hall.

APPENDIX 1A

Frequencies of content areas of test one in 2006 by question number and content area

part	Question number	Content areas					
		Reading comprehension	Vocabulary	Grammar	Pronunciation	Punctuation	Rm.
I	1						Ungrouped
	2						writing test
	3						writing test
	4						writing test
	5						writing test
	6						writing test
	7	√					
	8						writing test
	9						writing test
	10						writing test
II	11					√	
	12					√	
	13					√	
	14					√	
III	15						writing test
	16						writing test
	Total	1	-	-	-	4	10

APPENDIX 1B

Frequencies of content areas of final test in 2006 by question number and content area

Part	Question number	Content areas					
		Reading comprehension	Vocabulary	Grammar	Pronunciation	Punctuation	Rm
I	1		√				
	2						
	3			√			
	4			√			
	5			√			
	6			√			
	7		√				
	8			√			
	9			√			
	10			√			
	11			√			
	12			√			
	13		√				
	14		√				
	15			√			
	16			√			
	17			√			
	18			√			
	19			√			
	20	√					
II	21		√				
	22	√					
	23		√				
	24		√				
	25	√					
	26	√					
	27	√					
	28	√					
	29	√					
	30	√					
III	31 32 33 34						31-34 guided writing tests
IV	35						Writing test

	Total	8	7	14	-	-	5
--	-------	---	---	----	---	---	---

APPENDIX 2A

Frequencies of content areas of test one in 2007 by question number and content area

Part	Question number	Content areas					
		Reading comprehension	Vocabulary	Grammar	Pronunciation	Punctuation	Rm
I	1					√	
	2		√				
	3			√			
	4			√			
	5			√			
	6			√			
	7			√			
	8						
	9			√			
	10						Writing test
II	11		√				
	12		√				
	13		√				
	14		√				
	15		√				
	16		√				
	17		√				
	18		√				
	19		√				
	20		√				
	21					√	
	22					√	
	23					√	
	24					√	
	25					√	
III	26						Writing test
	Total	-	11	6	-	6	2

APPENDIX 2B

Frequencies of content areas of final test in 2007 by question number and content area

Part	Question number	Content areas					
		Reading comprehension	Vocabulary	Grammar	Pronunciation	Punctuation	Rm.
I	1						Writing test
	2						Writing test
II	1	√					
	2	√					
	3	√					
	4	√					
	5	√					
III	1		√				
	2		√				
	3		√				
	4		√				
	5		√				
IV	1			√			
	2			√			
	3			√			
	4			√			
	5			√			
	6			√			
	7			√			
	8			√			
	9			√			
	10			√			
	1						Writing test
	Total	5	5	10	-	-	3

APPENDIX 3A

Frequencies of content areas of test one in 2008 by question number and content area

Part	Question number	Content areas					
		Reading comprehension	Vocabulary	Grammar	Pronunciation	Punctuation	Rm.
I	1		√				
	2		√				
	3		√				
	4		√				
	5		√				
II	1		√				
	2		√				
	3		√				
III	1		√				
	2		√				
	3		√				
	4		√				
	5		√				
IV	1		√				
	2		√				
	3		√				
	4	√					
	5	√					
	6	√					
	7	√					
	Total	4	16	-	-	-	-

APPENDIX 3B

Frequencies of content areas of the final test in 2008 by question number and content area

Part	Question number	Content areas					
		Reading comprehension	Vocabulary	Grammar	Pronunciation	Punctuation	Rm.
I	1			√			
	2			√			
	3			√			
	4			√			
	5			√			
II	1		√				
	2		√				
	3		√				
	4		√				
III	1	√					
	2	√					
	3	√					
	4	√					
	5	√					
IV	1			√			
	2			√			
	3			√			
	4			√			
V	1					√	
	2					√	
	3					√	
	4					√	
	5					√	
VI	1		√				
	2		√				
	3		√				
	4		√				
	5		√				
VII	1						Writing test
	Total	5	9	9	-	5	1

APPENDIX 4A

TEST ONE

AWASSA COLLEGE OF HEALTH SCIENCES THE BASIC SCIENCES DEPARTMENT ENGLISH TEST ONE FOR THIRD YEAR STUDENTS -2006

I. CHOOSE THE BEST ANSWER FROM THE GIVEN ALTERNATIVES AND WRITER THE LETTER OF YOUR CHOICE IN THE SPACES PROVIDED

- _____ 1. One is different from others
A. speaking B. reading C listening D. none
- _____ 2. One is not a factor against which a content of certain piece of writing
could be evaluated
A. age B. relevance C. clarity D. brevity
- _____ 3. A camel is the ship of the lowland desert. The word which has been
repeated unnecessarily is
A. the B. ship C. lowland D. desert
- _____ 4. One is wrong about written language
A. Mistakes are note tolerated
B. difficult or impossible to get feedback
C. Possible to stop for further explanation
D. not a hint
- _____ 5. The error being observed in “Tesege lectured on planting and how
to harvest.” is
A. fragment C. faulty parallelism
B. dangling modifier D. none
- _____ 6. The observable problem with the sentences, “Scratching noises on
the roof indicated that I had a visitor.” and “Probably a tiger or a
wildcat.” is
A. Run on C. punctuation
B. choppiness D. Agreement

- _____ 7. Riding to the mountain, we saw a gang of robbers. Who was/were riding to a mountain?
- A. a gang of robbers C. mountain visitors
B. we D. possibly A&B
- _____ 8. What sentence error do you see in question number 7?
- A. faulty parallelism C. run on
B. dangling modifiers D. fragment
- _____ 9. Which one of the following paragraph components can be taken as contract between a writer and a reader
- A. Coherence C. transition markers
B. Unity D. none
- _____ 10. If a certain piece of writing ends as, “We can safely conclude and evidences clearly show that being member of small family is more advantageous over being member of a large family,” then such piece of writing is
- A. persuasive C. expository
B. narrative D. informative

PART II WRITE THE FOLLOWING SENTENCE WITH PROPER PUNCTUATION

1. Ato Bekele said the test was perfectly good

2. Dessie the capital city of Wollo is 300km from Addis Ababa

3. what a cheerful teacher

4. Senait went to supermarket and bought banana chocolate and fruits

PART THREE: WRITING

14. II. FIRST RE ARRANGE THE FOLLOWING JUMBLED SENTENCES INTO A LOGICAL ORDER AND THEN USE THEM TO WRITE A BRIEF PARAGRAPH.

- A. It causes lung cancer
- B. All these are bad
- C. smoking is very dangerous habit
- D. It also causes throat infection.
- E. It is, therefore, good to avoid smoking altogether.

Order: 1 _____ 2 _____ 3 _____ 4 _____ 5 _____

PARAGRAPH WRITING

15. _____

APPENDIX 4B

TEST TWO

**AWASSA COLLEGE OF HEALTH SCIENCES THE BASIC SCIENCES
DEPARTMENT ENGLISH FINAL EXAMINATION FOR THIRD YEAR
STUDENTS -2006**

**I. CHOOSES THE BEST ANSWER FROM THE GIVEN ALTERNATIVES AND WRITE
THE LATTER OF YOUR CHOICE IN THE SPACES PROVIDED.**

- _____1. Which one following compound nouns is mispelt?
A. Newspapers C. Mother tongue
B. Mother-tongue D. None
- _____2. Tsega: How old are you?
A. That is rude of you C. I will be 20 years of old
B. I was 20 Years age D. I am 20 years of age
- _____3. No matter how _____, Ethiopians are proud of themselves.
A. rich enough they may be C. Poor they are
B. Poor are they D. All
- _____4. The director hardly kikes wishy washy students who miss classes on
any flimsy grounds _____?
A. does he B. will he C. doesn't he D. all
- _____5. I am so pleased _____ that your request for marriage in the
coming year has got green light.
A. to condemn you C. Of telling you
B. to argue you D. to tell you
- _____6. Tadesse: Excuse me sir, Would I talk to Meseret?
Getachew: Sorry Madam, we have two Meseret's in the class. _____
Meseret: do you want to talk to?
A. what B. who C. Which D. all

- _____ 7. Patient: How shall I take this bottle of syrup?
Physician: Just read the _____ on the bottle; it tells you how to take to take it.
A. Direction B. Composition C. Ingredients D. None
- _____ 8. The Ambassador a long with the diplomats _____ our college.
A. have visited b. has visited C. has been visited D. all
- _____ 9. I as well as my friends _____ surprised when she won a 1000,000 birr lottery.
A. was B. were C. have d. has
- _____ 10. he ----- that is was his fault.
A. promised C. admitted
B. accused D. refused
- _____ 11. Christians ----- on Sundays.
A. is ever working C. never work
B. is never working D. never works
- _____ 12. ----- she worked hard, she was unable to finish her work.
A. However C. although
B. Because D. Despite
- _____ 13. We have to teach our children how to be thrifty beginning from their early ages. then they will be able to utilize the scarce resources at their reach very efficiently. The underlined word means.
A. healthy C. pround
B. jealousy D. economical
- _____ 14. Our farmers, nowadays, are able to produce food grains enough for domestic consumption as well as export.
A. local C. not wild
B. in the country D. A and B
- _____ 15. It is now ten years since he ----- to Ethiopia.
A. Wash coming B. Had coming C. came
D. Has come E. Did come

- _____ 16. If you borrow from bank, you will have to ----- the money with interest.
- A. Pay in B. Pay out C. Pay back D. Pay for
- _____ 17. The book was ---- that I forgot the time.
- A. Very interesting B. too interesting C. so interesting
- D. interesting enough E. Most interesting
- _____ 18. Your father works very hard, _____
- A. isn't it? B. doesn't he? C. Does he?
- D. won't he E. Shouldn't he?
- _____ 19. _____ People went to the cinema, while others went to the football match.
- A. Other B. Few C. All D. Some E. Plenty

II. READ THE PASSAGE BELOW CAREFULLY AND ANSWER THE QUESTIONS THAT FOLLOW

1. Men's contribution to house work and child care, which we may call reproductive work, is negligible compared to women's contribution towards productive, or income-earning, work. In countries like Tanzania, Burkina Faso and Java men spent a matter of minutes each day on domestic work. It is men's failure to share the domestic burden that accounts for women's longer working days.

2. To understand this we need to take a closer look at this domestic burden. We have all heard of the long hours it takes to grind grain, collect fuel wood and fetch water, wash clothes, prepare and cook food, when there is no electricity, piped water or washing- machines. How can women possibly manage all this work in only five or so hours per day?

3. The answer is that children also work hard. By the time they reach their teens, children are generally providing at least half as many hours of work as adults despite the girls, is in time-consuming chores like collecting fire wood and fetching water, and 4 looking after their younger brothers and sisters. Indeed if they didn't help with these tasks, 5 the burden of reproductive work would be truly impossible for adult women to bear. In the Javanese village, for example, the total amount of reproductive work needed each day to maintain a household was nearly 10 hours. It would be almost impossible for them to do anything else. Adult women are able to spend time working on productive or income-earning activities only because their daughters-and, to a lesser extent sons, and to a still lesser extent, husbands share part of the reproductive work.
4. Of course, this understanding of how domestic workload is managed does not provide any answer by it self, but at least it emphasizes the dilemmas. It exposes the ignorance-and arrogance involved in the assumption made by many development agencies that rural women's time is 'free' time, ready to be filled by development activities. But it also highlights the dangers of the opposite assumption: that women are so burdened already with domestic work that projects should concentrate on lightening this burden, leaving the 'real' work of income generation to others. this is because access to an independent income is an important source of security and status to women.

Adapted from New Internationalist, 1988 with slight modification.

- _____21. Negligible line (2) means
- | | |
|-----------------|------------------|
| A. Considerable | C. Insignificant |
| B. Significant | D. All |
- _____22. Productive work means

- A. Hard work C. Industrious
- B. diligent D. Income-earning

_____ 23. Dilemmas means

- A. Problems C. Clarity
- B. Interest D. All

_____ 24. Chores means

- A. Persons who work tighter C. Activities
- B. Chants D. All

_____ 25. Which one is wrong according to the passage?

- A. Men's involvement in reproductive work is greater than that of women
- B. Although domestic chores take longer hours, they are not to be considered as burden.
- C. Teenagers work for twice as many hours as adults.
- D. All

_____ 26. One is true according to paragraph one

- A. Men's contribution to productive work is greater than that of women
- B. Men in Tanzania, Burkina Faso and Java are admired for their contribution to works house work and child care
- C. Men failed to take part in reproductive chores so that women are forced to bear high burden
- D. A and C

_____ 27. The main idea that paragraph two makes is

- A. Women take longer hours to do minute things.
- B. What women do at the home take long hours for the absence of technologies and that contributes to women's longer working days.
- C. Women manage to accomplish their duties for they are assisted by their children.
- D. All

_____ 28. 'which' in to house work and child care, which we may refers to

- A. Men's contributions C. women's contribution
B. Productive work D. House work and child care

_____ 30. 'them' In '..... almost impossible for them to do ' refers to

- A. boys B. teenagers C. girls D. women

WRITING

35. III. RE-ARRANGE SET OF ITEMS (31-34) TO MAKE A SENSIBLE SENTENCE.

NOTE THAT YOU WILL ADD NOTING TO ITEMS GIVEN

31. The /up/ straight/pick/teacher/and/sit/your pens/said

Ans: _____

32. The boss/all the subordinates/agreed with/said/what

Ans: _____

33. The physician asked me, "What is wrong with you?" Report this (later)

Ans: _____

34. The teacher reminded us not to write more than one page. Convert into quote.

Ans: _____

IV. WRITE A BRIEF PARAGRAPH USING THE INFORMATION GIVEN BELOW.

Aster

- Was born in shashemene
- Enrolled in a school at the age of 8
- Attended high school in shashemene
- Failed ESLCe but joined Technical

Nigist

- Was born in Yirgalem
- Began school at the age of 8
- Completed secondary education in Awassa
- Passed ESLCE with distinction,

and Vocational Training Institute
and got 10+ certificate in Accounting

joined AAU and specialized in
Political Science and
international Relations

APPENDIX 5A

TEST ONE

AWASSA COLLEGE OF HEALTH SCIENCES THE BASIC SCIENCES DEPARTMENT ENGLISH TEST ONE FOR THIRD YEAR STUDENTS -2007

I. CHOOSES THE BEST ANSWER FROM THE ALTERNATIVES GIVEN AND WRITE THE LETTER OF YOUR CHOICE IN THE SPACES PROVIDED.

- _____ 1. Which one is correctly punctuated?
- A. when they arrive, what shall I tell them to do?
 - B. When they arrive? What shall I tell them to do?
 - C. When they arrive. What shall I tell them to do?
 - D. When they arrive what shall I tell them to do?
- _____ 2. A number of _____ were seen watching the games of the 26th African cup of nations.
- A. Referee
 - B. Teams
 - C. spectators
 - D. Players
- _____ 3. The book was _____ that I forgot the time of appointment.
- A. Interesting enough
 - B. So interesting
 - C. too interesting
 - D. such interesting
- _____ 4. She _____ if she had been able to.
- A. Would visit
 - B. would be visiting
 - C. will have visited
 - D. would have visited
- _____ 5. Aster was _____ her new traditional clothes when I met her in the gathering.
- A. Putting on
 - B. Wearing
 - C. having on
 - D. all
- _____ 6. _____ you say, you will not persuade a rigid person.
- A. However
 - B. Despite
 - C. whatever
 - D. Although

- _____ 7. A: What do you prefer?
 B: _____.
- A. I always prefer tea to coffee.
 - B. I prefer always tea than coffee
 - C. I prefer either tea or coffee
 - D. All
- _____ 8. All of the following points should be considered when one conducts public speech except
- A. Organize the main points and rehearse
 - B. Be formal both at the beginning and ending speech
 - C. Have eye contact with the audience
 - D. Be fearful and shy to certain extent.
- _____ 9. Which one of the following is the functional classification of a sentence?
- A. compound sentence
 - B. Interrogative sentence
 - C. Simple sentence
 - D. Complex sentence
- _____ 10. One qualifies both note and prose summary
- A. Relevance
 - B. Brevity
 - C. clarity
 - D. all of them

VOCABULARY

MATCHING ITEMS

II. COMPLETE THE BLANK SPACES IN THE SENTENCES IN COLUMN A WITH THE WORDS FROM COLUMN B

A

- _____ 11. Did they_ safely in Gambella
- _____ 12. Her father_____ her to go.
- _____ 13. It's the biggest lie he had_____ told
- _____ 14. They ____ that they had taken the money
- _____ 15. In Addis Ababa it _____ rains in August
- _____ 16. Be careful! Thieves may _____ your money

B

- a. reach
- b. ever
- c. denied
- d. steal
- e. steel
- f. aloud

_____ 17. I'm _____ waiting for them to arrive.

g. still

_____ 18. tell me all _____ you know about it

h. that

_____ 19. Do you know _____ it looks like

i. always

_____ 20. The teacher wants to _____ about my work

j. what

k. discuss

i. forbade

m. talk

n. rob

o. arrive

III. PUNCTUATE AND RE-WRITE THE FOLLOWING SENTENCES

21. Ato Samuel said that the exam was quire good

22. Where are you going

23. Awassa, which is the capital city of SNNPRS is 20 kms from shashemence

24. What a beautiful lady

25. When shall we get married said wudinessh

26. IV. WRITE A BRIEF ARGUMENTATIVE PARAGRAPH USING ONE OF THE FOLLOWING TOPICS

a. HIV/AIDS is accurse from God

b. Gender equity is not a necessity

c. Corruption must not be condemned in the poor countries like Ethiopia.

d. Affirmative action is meant for girls only

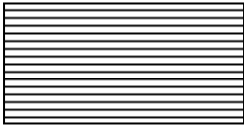
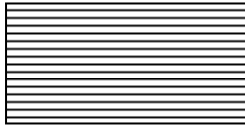
APPENDIX 5B

TEST TWO

**AWASSA COLLEGE OF HEALTH SCIENCES THE BASIC SCIENCES
DEPARTMENT ENGLISH FINAL EXAMINATION FOR THIRD YEAR
STUDENTS -2007**

PART I

**THIS IS AN EXAMPLE OF BUSINESS COVER LETTER. SHOW THE CORRECT
STRUCTURES FOR THE NUMBER GIVEN ON SAMPLE LETTER.**

2  3. _____ 4  7 _____	1  5. _____ 6. _____
---	--

Answer Here

1. _____
2. _____
3. _____
4. _____
5. _____
6. _____
7. _____

WRITE AT LEAST THREE IMPORTANT THINGS WHEN WE ARE WRITING AN APPLICATION LETTER

1. _____
2. _____
3. _____

CASUALTY

While John smith was going to school this morning, a car knocked him down. His right leg was broken just below the knee. Some people who saw the accident laid him in a comfortable position on the pavement and telephoned for an ambulance. John was badly shocked and in great pain. while the people were waiting for the ambulance, the driver of the car that knocked john down covered him with a coat and tried to comfort him.

When the ambulance arrived, John was lifted onto a stretcher, put into the ambulance and driven to a hospital that was not far away. When he arrived at the casualty department, he was admitted. The nurse who admitted him gave him an injection of morphine to combat the shock and the pain. The doctor who examined him comforted him and told him everything would be all right. While the doctor was examining him, the nurse who had admitted him took his pulse and blood pressure. No anti-tetanus injection was necessary as john had been immunized against tetanus six months earlier.

As John's wound was lacerated and contuse, he was given an antibiotic to prevent the onset of infection. He was then taken to the x-ray department. While the x-ray examination was being carried out, John's parents arrived at the hospital. The doctor who examined John told them that john had sustained a compound fracture of the tibia the fibula. He explained that an operation was necessary and asked Mrs. Smith to sign consent for operation form. Mrs. Smith signed the consent form while Mrs. Smith was giving John's personal particulars and previous medical history to the nurse. Before

John's parents left the hospital, they were told when they could visit him and were given a list of things that he would need.

While John was waiting to go to theatre, a quarter-hourly record of his pulse and blood pressure was kept and he was given a suitable premeditation.

In the theater, John will be given an anesthetic and his wound will be explored. The x-rays which were taken by the radiographer will show the surgeon the exact site and extent of the fracture. Dead or dirty tissue will be excised and any small splinters of bone will be removed. The surgeon will perform the operation using a "no touch" technique to reduce the chances of infection.

When the surgeon is satisfied that the wound is as clean as possible, he will dust it off with an antibiotic powder. He will then suture it and cover it with a sterile gauze dressing. any areas that are raw will be dressed with sterile Vaseline petroleum jelly gauze. The surgeon will now be able to treat the fracture as closed. He will reduce that bones in to their anatomical position, and the leg will be put in plaster of Paris. John will then be taken back to an orthopedic ward.

When a post-operative patient like John arrives in the ward, he is carefully lifted from the stretcher and laid in a specially prepared bed. The foot of the bed is often raised on wooden blocks. The patient is placed on his back with his head turned to one side. As he is generally still unconscious, an artificial airway is in position and a nurse stays with him to ensure that his airway is kept clear. She also checks his pulse rate and volume, his colour and the rate and depth of his respirations.

Soon after John had returned to the ward, he began to regain consciousness. The nurse who was looking after him removed the airway from his mouth and gave him a pillow for his head. For the next few hours he slept soundly.

From time to time the toes of his injured leg were examined to see if they were warm and pink and his pulse and blood pressure were taken half-hourly.

At 6 P.m. John woke up and complained of severe pain in his leg. The surgeon who had performed the operation had prescribed pethidine if John complained of pain, and he was given an intramuscular injection of 50 mgs of pethidine at 6.10 p.m. As his blood pressure was now within normal limits, the bed-blocks were taken away and a bed cradle was put in his bed to take the weight of the bedclothes off his legs. A nurse offered him a bottle, but he said he could not manage to pass water.

A houseman visited John during the evening to check that he was all right and that he would be able to sleep. He prescribed a second injection of Pethidine, which was to be given at midnight if John complained of further pain. Two nurses came and helped John to wash his hands and face and to change from the white theatre gown into his own pajama jacket. John, who had been allowed frequent sips of water because he had not complained of nausea, was now given a cup of tea and told nurses that he was beginning to feel fine.

PART II

QUESTIONS FROM THE PASSAGE

_____1. Why wasn't John given an anti-tetanus injection?

- A. to compact the shock and the pain
- B. Because he had been immunized against tetanus all in his life
- C. I since he had been immunized against tetanus six months earlier.
- D. For the reason that he has refused

_____2. What was Mrs. Smith asked to sing?

- A. Operation consent
- B. A list of things he has

- C. permission to visit John
- D. None

_____ 3. According to the passage who has confirmed the surgeon the exact site and extent of the fracture.

- A. surgical doctor
- B. Radiographer
- C. John's Family
- D. The nurse

_____ 4. Why do you think the surgeon use a 'no touch' technique?

- A. to reduce pain
- B. to increase the chances of pain
- C. to decrease the possibility of infection
- D. All are wrong

_____ 5. All the time, why does the nurse remain with him?

- A. to ensure his air way is kept clear
- B. to check his pulse rate and volume, his colour and the rate
- C. To check the depth of his respiration
- D. All are the correct answer

PART: III

MATCHING

THE FOLLOWING LIST OF WORDS ARE TAKEN FROM THE PASSAGE SO THAT MATCH THE CONTEXTUAL MEANINGS OF WORDS IN COLUMN "A" FROM COLUMN "B"

A.

B.

_____ 1. Excised (para-4)

A. A thin cotton fabric used for covering and protecting wounds.

_____ 2. Splinter (Para-4)

B. To remove something completely

_____ 3. Dust off (Para-5)

C. To remove dust dirt etc from something

_____ 4. Suture (Para-5)

D. thin sharp Pieces of bone that has broken off a larger piece

_____ 5. Gauze (Para-5)

E. To sew up a wound

PART IV

GRAMMAR

I. CHOOSE THE CORRECT ANSWER TO COMPLETE THE SENTENCE

- _____ 1. He _____ when his mother came to wake him up.
A. slept B. was sleeping C. have slept D. is slept
- _____ 2. Belayneh _____ to his office nearly every day
A. walks B. walk C. is walking D. has walked
- _____ 3. Mrs. Smith _____ an injection every four hours
A. is given B. were given C. Are giving D. have been given
- _____ 4. Food poisoning _____ by bacterial
A. has caused B. is caused C. Is causing D. had caused
- _____ 5. She was watching TVE when She _____ the bad news.
A. heard B. hear C. had heard D. was hearing
- _____ 6. They _____ for war on large scale since last year.
A. prepare B. had prepared
B. have prepared D. Are preparing
- _____ 7. Before levis poster died he _____ the new science of bacteriology
A. should have found C. has found
B. was found D. had found
- _____ 8. The meeting _____ on next Monday
A. will held b. will be held C. is hold D. has hold
- _____ 9. I _____ a letter since early morning
A. wrote B. write C. have written D. will write
- _____ 10. These patients _____ yesterday morning
A. has catheterized C. have catheterized
B. were catheterized D. was catheterized

PART VI

THE FOLLOWING SENTENCES ARE NOT ARRANGED IN THE CHOROLOGICAL ORDER. THERE FOR INDICATED THE CORRECT SEQUENCES LOGICALLY ON THE SPACE PROVIDED

1. Because of this, they were given very little attention by scientists and no attempt was made to control them.
2. Many years ago, mosquitoes were regarded as insects with a nasty bite but of no more significance
3. In short, while, we have won half of the battle with mosquito, much more remains to be done to completely eradicate or control malaria.
4. Eventually, with the discovery of their role in the spread of malaria and yellow fever, they became a major focus of scientific research
5. For example, thousands of people in the lowlands of Ethiopia die of malaria
6. As a result, yellow fever has been eradicated, and is no longer a treat to human life.
7. In contrast, malaria remains un conquered in spire of the human effort to control.

Correct order _____

APPENDIX 6A

TEST ONE

AWASSA COLLEGE OF HEALTH SCIENCES THE BASIC SCIENCES DEPARTMENT ENGLISH TEST ONE FOR THIRD YEAR STUDENTS -2008

PART I. CHOOSE THE WORD THAT SUITS FOR THE BLANK SPACES PROVIDED BELOW

Training, domestic, Official, effect, supposed to, share frighten, typical, affect

- _____ 1. Horror films _____ children, and should be banned by law.
- _____ 2. Economic _____ policy should be adopted and implemented as soon as possible.
- _____ 3. Helping other people is the _____ behavior of some good people
- _____ 4. Fluctuation of weather can _____ people's lives adversely
- _____ 5. _____ trade does not incorporate import-export business.

II. WRITE THE ROOT OR BASE WORD FROM EACH WORD GIVEN BELOW

1. Maladjustment
2. irresponsibility
3. Unachievable

III. GUESS THE MEANING OF WORDS UNDERLINED IN EACH SENTENCE IN DISCIPLINED WAY.

1. The president has competed his period of presidency. He is also at the age of retirement, unless we back him up, he will not be reelected to preside.

Back him up → _____

2. We haven't been met each other for ages. Will you be at home, so that we may call on you tomorrow? Call on → _____

3. If you call me up at any time tomorrow, you will know where I am.

Set off → _____

4. The students set off early to reach the school on time.

Set off → _____

5. You need to join in the discussion in order to workout the problems given by the instructor.

Join in → _____

IV. READ THE FOLLOWING PARAGRAPH AND ANSWER THE QUESTIONS ACCORDINGLY

In many societies, the man is still seen as the one who provides for and protects the family. But he can and should help with raising the children in other ways as well. The parents should make decisions together about caring for the children. For example, when to take them to a health facility, and when they should start school. He can also help counsel and advise them as teenagers, discussing issues like when to get married, and what career they should train for. The man and other family members can share the task of watching the children and should help teach the children how to cope with responsibility as they grow up. In particular, the father should ensure that his daughters are given equal opportunities as his sons in terms of education, healthcare, and other benefits.

1. "Teenagers" on line 6 refers to _____?

- a. Adolescents aged above 20
- b. Persons aged between 10 and 20
- c. Parents who take care of their children

2. "Career" on line 7 means?

- a. Children who preserve pathogenic organism in their body.
- b. Job or skill for their future life
- c. Parents who suffer for their children

3. To cope with on line 9 means____?

- a. Shoulder or become ablest to
 - b. fight and over come every walk of life
 - c. to be nervous and energetic
 - d. none
4. The writer blames the societies on one hand and provides recommendation on how to act or perceive on another hands.
- a. What dos he blame?
 - b. what are recommendations he provides to both societies, and to the father?
 - c. give appropriate title for the paragraph above
 - d. writer the idea of the paragraph at most in two lines (<50 words)

APPENDIX 6B

TEST ONE

**AWASSA COLLEGE OF HEALTH SCIENCES THE BASIC SCIENCES
DEPARTMENT ENGLISH FINAL EXAMINATION FOR THIRD YEAR
STUDENTS -2008**

I. JOIN THE FOLLOWING SENTENCES BY WORDS BELOW

But	And	Because	Who	Either	Neither
-----	-----	---------	-----	--------	---------

1. My brother came _____ we went for a walk.
2. They tried hard, _____ they didn't succeed.
3. I saw the man _____ was blind.
4. She forgave him _____ he asked her for apology
5. Alemu _____ returned the coat nor paid the price

II GUESS THE MEANING OF WORDS USED IN CONTEXT GIVEN BELOW

1. Ministry of health works in one way or another, to reduce the morbidity and mortality rate in the respective country.
2. Factors that intensify the risk condition should be identified ahead of time and taken measure in time.
3. Primary hypertension is most often asymptomatic until sever injury has began to develop.
4. Hypertensive patients, if they are smokers, should be advised to quit because it aggravates the condition indirectly.

Morbidity _____

Mortality _____

Intensify _____

Asymptomatic _____

Auit _____

III. COMPREHENSION QUESTION

READ THE FOLLOWING PASSAGE AND ANSWER QUESTIONS GIVEN BELOW

1. Chronic hypertension is associated with increased morbidity and mortality. Left untreated, chronically elevated blood pressure can lead to heart disease, kidney disease, blindness and stork. The degree of injury is directly related with the degree of the blood pressure elevation. In addition, risk factors like obesity and sedentary life style intensify the injury. Deaths related to chronic hypertension is largely resulted from cerebral hemorrhage, heart failure, renal failure and myocardial infraction. Despite its potential for serious harm, hypertension usually remains asymptomatic until long after injury has began to develop.
2. The treatment of hypertension has two objectives: reduction of blood pressure to or < 140/90 mmHg and prevention of long term complication.
3. Life style changes termed as non pharmacological therapy can decrease the blood pressure of persons with primary hypertension whose cause is not identifiable. As there is direct relation between obesity and blood pressure, weight reduction has some thing to do with it. Thus, dietary restriction, sodium (salt) and alcohol restriction along with regular aerobic exercise are recommended to reduce blood pressure. The victims are also encouraged to quit smoking for it has indirect aggravating effect.

1. Write three titles for three paragraphs (one for each).

Para. 1 _____

Para. 2 _____

para.3. _____

2. Write three sentences that summarize the passage.

Para 1. _____

Para 2. _____

Para3. _____

3. Write two basic objectives of treating hypertension.

1. _____

2. _____

4. Deaths from hypertension can largely be resulted from _____

5. As it is given in the passage above, what are non pharmacological therapies of hypertension?

IV CHOOSE THE BEST ANSWER FOR THE FOLLOWING QUESTIONS

_____ 1. Which one of the following sentences is structurally complex?

A. Tesfaye built the house, but he sold his car

B. He who gives a help to the need saves good treasure to him self.

C. Hiwot will build a house or buy a car

D. all can be possible answer

_____ 2. Molla washed his clothes; _____, he didn't take a bath. The possible word to be filled in the blank space is _____

A. more over B. however C. hence D. all can be possible words

_____ 3. The president together with all high authorities _____ been visiting hewn churches of Lalibela.

A. has B. have C. is D. all can be possible answers

_____ 4. Neitehr neither the patient nor the doctors mistaken concerning the dispute between them.

A. or B. but C. however D. have E. all can be the answer

V. PUNCTUATE THE FOLLOWING SENTENCES CORRECTLY

1. Nelson Mandella who later become the president of south African spent many years in the prison

2. The book we re studying is an introductory to biology

3. Look at the kitten its wagging the tail

4. Our English teacher Ato Alemu used to live in Awassa

5. What are you doing there up to this time the girl asked him angrily

VI. MATCH THE FOLLOWING VOCABULARIES WITH THEIR EXPLANATION

A

B

_____ 1. Draw

a. time of human development before the invention writing

_____ 2. Amateur

b. the time after invention of writing

_____ 3. Tour

c. the result of the game in which both sides have different scores

_____ 4. Accommodation

d. the result of the game in which both sides have equal score

_____ 5. Pre historic time

e. Person who plays the sport game with out payment.

f. A general term for a place to stay

VII. WRITING

DEVELOP THE FOLLOWING TOPIC SENTENCE INTO A PARAGRAPH.

There are a lot of materials to read in the library.

DECLARATION

I, undersigned, declare that this thesis is my original work and has not been presented for a degree in any other university.

NAME: ASMARE MEHRET MENGISTU

SIGNATURE: _____

PLACE: AAU. INSTITUTE OF LANGUAGE STUDIES

SUBMISSION DATE: JULY, 14/2008