

AN EVALUATION OF ENGLISH LANGUAGE FINAL EXAMS: A CASE OF DEBRE MARKOS PREPARATORY SCHOOL

Dawit Dibekulu Alem

Research Scholar, Department of English language Literature, College of Social Sciences and Humanities, Mekdela Amba University, Ethiopia

Received: 02 May 2021

Accepted: 07 May 2021

Published: 31 May 2021

ABSTRACT

The aim of this paper is to evaluate the English language final exam based on scholars view point. In doing so descriptive design was employed and qualitative and quantitative data analysis was employed to analyses the data obtained from the document. Document was the primary data for this study. So, the findings indicate that the test was not prepared in effective way by considering difficulty level, test variety, skill variety. Thus, teacher in school need to prepare the test by considering the scientific method of preparing an exam.

KEYWORDS: *Evaluation, Test, Language*

INTRODUCTION

Language testing is the essential and main part of teaching language. It has passed a numbers of years, so due to the advancement of different theories and approaches in language teaching language testing also developed in terms of the this developments.

The test needs to be effective, efficient, reliable and valid to achieve the goal of language teaching and learning. So, the aim of this paper is to evaluate DebreMarkos preparatory school grade 11th English second semester final exam which was given in 2011 E.C. And it incorporates: what is test why we evaluate test?, background of the test, evaluation of the test based on the criteria, external evaluation, evaluation on the coverage of the contents, coverage of the test all the contents in the book, does the test measure all, level of the test difficulty and easiness, level of the test difficulty and easiness in terms of the questions, level of the test difficulty and easiness in terms each question, validity and reliability of the test, biasness of the test, evaluation of the instructions, quality of the choices/ distracters, grading or marking, contextually, limitedness, test authenticity and instructiveness, editing, economy of the exam, and finally, conclusion remarks were drive.

THEORETICAL OVERVIEW

According to Madsen, testing helps not only the students but also to teachers. He mentions two ways how the well-made tests can help learners. First, he says, such tests can help create positive attitudes in terms of motivation and efficient instruction. This means that a sense of accomplishment should be taken into account. Madsen believes that tests of appropriate difficulty announced well in advance and covering skills scheduled to be evaluated, can also contribute to a positive tone by demonstrating your spirit of fair play and consistency with course objectives (Madsen 1983:4). The second way that students can benefit from tests is by helping them to master the language. They can confirm what each person has

mastered and they point up those language items that need to be studied further. It can help learners to adjust their own personal goals (Madsen 1983:4).

Some scholars stated about the reason for test by asking the question: Why test? So (Heaton 1990; Madsen 1983) stated that we test to:

- Diagnose students strengths and needs
- Provide feedback on student learning
- Provide a basis for instructional placement
- Inform and guide instruction
- Communicate learning expectations
- Motivate and focus students' attention and effort
- Provide practice applying knowledge and skills

As mentioned earlier, testing also helps teachers. Due to testing they can be able to answer the important questions, such as:

- Have I been effective in my teaching?
- Are my lessons on the right level?
- Am I aiming my instruction too low or too high?
- Am I teaching some skills effectively but others less effectively,
- What areas need more work?
- Which points need reviewing?
- Should I spend more (or less) time on this material with next year's students?
- Were the test instructions clear?
- Was everyone able to finish in the allotted time?
- Did the test results reflect accurately how my students have been responding in class and in their assigned work?
(Madsen,1983:5)

In other words, testing can be used to diagnose both teachers' and students' effort. It can confirm progress that has been made and show how to redirect our future efforts. Madsen adds that good tests can sustain or enhance class morale and aid learning (Madsen, 1983).

Heaton presents reasons for testing as follows: finding out about progress, encouraging student, finding out about learning difficulties, finding out about achievement, placing students, selecting students, and finding out about proficiency (Heaton, 1990:9–18)

METHODOLOGY

Research Design

As mentioned earlier, this research project was aimed to evaluate the English language final exam based on scholars view point. Descriptive design was employed.

Data Collection Instruments

Document analysis was an instrument that permits a level of in-depth information and since the current study is aimed to evaluate English language final exam document analysis is the best suited data collection instrument.

Methods of Data Analysis

The data collected were analyzed using both qualitative and quantitative method. Then, based on the findings, conclusion and recommendations were made.

DATA ANALYSIS AND DISCUSSIONS

Evaluation of the Exam Based on the Criteria

- **External Evaluation**

Since very exam paper need to have external part/ cover page this part need to include the space to write students personal information, time, number of questions, exam types, subject etc. so, We tried to Evaluate external part of DebreMarkos preparatory school grade 11 second semester English final exam as follow:

Strong Sides:

- The exam cover state name of the school, subject, grade level, time allowed.

Weak Sides

- Even though there is an answer sheet for the exam with space provided to fill students' personal profile the exam must include space on the cover page to reduce ambiguity.
- The exam load, number of questions, and parts of the exam need to be stated on the cover page.
- Even though the exam question were written in two sides of the paper to save it this will lead the students miss on part so it is good if the exam was written on one side of the paper.
- The exam should be better if there is a general direction on the cover page (like restriction rules).

Background of the Test

In this part of our paper we are intended to explain the general explanation of the exam as a whole. This exam was prepared in Amhara Regional state East Gojjam Zone DebreMarkos Town DebreMarkos Preparatory School for grade 11th students for second semester. The exam has a total of forty (40) questions. And to see the questions in terms of the content of preparation and their item types let us look the following table:

Table 1 shows we can understand that the exam includes all the skills in the language except listening and writing. But in language teaching all skills need to get equal attention. Even though it difficult to test listening skills in the final exam the teacher can have a chance to test writing skills by controlled and guided writing activities. On the other hand in

terms of the question the exam has only multiple choice questions this makes the exam is too poor. We think that the exam is prepared for the advantage of teacher to save time for marking it. But it is not good to measure students' performance / intelligence. In relation our idea, Madsen (1983) stated that testing can be used to diagnose both teachers' and students' effort, It can confirm progress that has been made and show how to redirect our future efforts, and good tests can sustain or enhance class morale and aid learning. So, this can be justified by incorporating all the contents in the book the book also covers all skills and the teacher teach all the contents which able to teach the language skills but in this we did not see such things rather it focus only three skills: grammar, vocabulary, reading and speaking but two skills (listening and writing) are neglected. To supplement all skills the items which incorporated also has a significant role to test all skills but since this exam includes multiple choice items it fails to test the language skills. Furthermore, the text book includes so many linguistic issues for the benefit of achieving linguistic competence of the students but as we have seen this exam it losses much contents in the book even though it tried to cover some parts.

So, as a teacher we can recommended that even though it is difficult to test all the contents in the final exam it is better if it includes some basic important points need to be test tested in doing this we can assure the content validity of the test. On the other hand in terms of skills as language test the exam should be prepared by considering all skills but not only in this exam all other exams leave listening and little bit writing skills so as an English language exam it should be prepared by considering all the skills of the language and also to become valid in contents it should be good is it concentrate this things.

Table 1: Skills and Items in the Exam

No	Skills	Items	True / False	Matching	Multiple Item	Blank Space	Short Answer	Essay	Percentages
1	Reading	0		0	10	0	0	0	25 %
2	Writing	0		0	0	0	0	0	0 %
3	Listening	0		0	0	0	0	0	0 %
4	Speaking	0		0	5	0	0	0	12.5 %
5	Grammar	0		0	15	0	0	0	37.5 %
6	Word order	0		0	3	0	0	0	7.5 %
6	Vocabulary	0		0	7	0	0	0	17.5 %
Total		0		0	40	0	0	0	100 %
Percentages		0		0	100%	0	0	0	100 %

EVALUATION ON THE COVERAGE OF THE CONTENTS AND MEASURE OF INTELLIGENCE

Coverage of the Test All the Contents in the Book

Since the ultimate goal of a certain test especially final exams is to check whether the students understand the entire lesson they learnt in the teaching learning process the exam need try to cover all the contents in the book(the contents in the semester). So, we try to see this exam in terms of second semester contents from chapter seven to twelve.

As it is shown in Table 2 the exam covers more grammar questions (which consists 12 question) 30 % of the exam but as we observed in the book we observed so many things related to grammar but in terms of the content the exam is not covered all the contents, and secondly 10 question (25 %) percent of the question were reading comprehension so this exam includes this skills sufficiently, thirdly, 10 questions (25 %) of the questions were readings questions were speaking skills, 2 questions were conjunctions and clauses, 1 question were related to articles and prepositions but this is short and can't measure students intelligence in the area. But in the book there are so many topics are includes but this exam does not cover the topics sufficiently. So, as per final exam level the exam should cover at least the most important

points need to be test. These things make the exam more effective and measure students' level of understanding in the contents of the book.

Table 2: Content coverage of the Exam

No	Contents which are Incorporated in the Book (Chapter Seven to Twelve)	No Question to Measure the Contents	Percentage
1	Reading comprehension	10	25 %
2	Listening	0	0
4	Affixes: Prefixes and suffixes	0	0
5	Speaking: Dialogue, pronunciation, role play, and discussion	5	12.5 %
6	Punctuations	0	0
7	Conjunctions, clauses,	2	5 %
8	Parts of speech	0	0
9	Grammar: tenses, phrasal verbs, and conditionals	12	30 %
10	Writing :essay, summary, and reports	0	0
12	vocabulary : word power, completion, word building, and Contextual meanings and direct meaning of words	10	25 %
13	Article and prepositions	1	2.5 %
Total		100	100 %

Does the Test Measure All Intelligences?

If we raise this issue we need to raise different issues in line with this. Mainly, one educational curriculum design with consideration of three things like educational domains (cognitive, affective, and psycho motor domain) the easy part we evaluate the cognitive domain but we also measure in other items so, in this exam the all the question are measured the cognitive domain but to some extent affective domain also include especially in communicative part but psychomotor domain is not that much measure in the exam. There are different levels of questions. These are: Knowledge Question (are questions about knowledge. Instead of focusing on specific content, they focus on how knowledge is constructed and evaluated, and it represents the lowest level of learning out comes in the cognitive domains, Comprehension Questions (This may be shown by translating material from one form to another (words or number), by interpreting material (explaining or summarizing and by estimating future trends (predicting consequences or effects), Application Questions(refers to the ability to use learned material in new and concrete situations), Analysis Questions (Analysis refers to the ability to break down material into its parts so that its organizational structure may be understood), Synthesis Question (refers to the ability to put parts together to form a new whole. synthesis can pull together many disorganized elements or parts so as to form a whole.), Evaluation Questions (Evaluation is the ability to judge the value of material (statement, novel, poem, research report) for a given purpose) So, we tried to see the question in intelligences level as follows:

As in Table 3 shown that the test tries to cover the level of intelligence but we observe some gaps because the exam tends to measures comprehension ability of students (75 %) and little bit synthesis(7.5 %) and remembering (17.5 %) questions but the left three intelligences level. But in this exam the three level of intelligence are missed to measured and checked students ability of analysis, application and evaluation the question related to this thing need to be included. By doing this thing we can measure students intelligence. So, this indicates that the exam measures half of the intelligence level of the student but the as grade 11thstudents' they need to be assed all the skills and knowledge expected from them.

Table 3: Level of Intelligence and Number of Questions

No	Level of Intelligences	Incorporated Questions	Number of Questions	Percentages
1	Remembering questions	1,2,3,4,7,8,9	7	17.5 %
2	Comprehension question	5,6,10,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,40,	30	75 %
3	Application question	-	-	-
4	Analysis question	-	-	-
5	Synthesis question	18,19,20,	3	7.5 %
6	Evaluation question	-	-	-
Total		40	40	100 %

Evaluation of the Exam in Terms of Test Types

We also see this exam in terms of Madsen presents following classification test as follow:

In Table 4 we tried to see the exam in terms of Madsens' classification of test types it indicates that the exam in terms of measuring type 45 % of the questions measured their knowledge of grammar and other skills but the 55 % measures their skills like inference, interpretative, speaking, vocabulary and others. So, in terms of this the exam is proportional. While in the second case in terms of students' response the exam is 100 % objective because it is multiple choice there are no subjective questions in this exam due to this it may not able to measure their analysis and other skills needed to be measured by subjective questions. And in terms of skill measurement receptive skill questions are 87.5 % and are 12.5 % of the questions are productive skills. So from this we can understand that this exam is tends to biased s in measuring only receptive skill.

Table 4: Taken From Contrasting Categories of ESL Tests (Madsen 1983:8)

Test Types	No of Questions	Percentages
Measuring		
Knowledge tests	18	45 %
Performance (or Skills) tests	22	55 %
In Terms of Answering		
Subjective tests	-	-
Objective tests	40	100 %
Skills		
Productive tests	5	12.5 %
Receptive tests	35	87.5 %

LEVEL OF THE TEST DIFFICULTY AND EASINESS

Level of the Test Difficulty and Easiness in Terms of the Exam

As stated in the above in order to help the students to perform easily the question the exam should be presented simple to complex questions: so, the students will do effectively perform the exam effectively and motive to do it in effective manner. Even though this exam try to order in such a way but in language testing the order should be from word to essay par but in this exam the passage comes so the student will be tired off to do other easy question but easy question may serve as a motive to do the exam effectively. If the question were order by considering their difficulty level it should be more effective than the previous on. As we try to glance the exam in general the exam is not that much related to the ministry of education exam but grade 12 student need to be assed in such away because if they are tested by the same level with grade 12 entrance exams.

Level of the Test Difficulty and Easiness in Terms of Each Question

Under this part we try to see the each question especially the question which has a serious problem in detail. In language testing the exam should be incorporate ever contents in the book which is relevant to students' intelligences. In reading comprehension part the question are good and fit with students' capacity. Question number, one, two and three measure text understand level of students while four and nine also inferring the other also contextual meaning so the questions are good. In similar way the grammar part is also fit with students' capacity but we have a doubt on the easiness of the question as preparatory student the level of difficulty of the question as equal as the iterance exam. So, the grammar part is not that much difficult or measure students capacity. But we see different question like question number seven the students can answer easily by scanning the paragraph but this question will be better if it was asked atclassroom assessment rather than at final exam level.

Reliability of the Test

In language testing reliability of the test need to be measured. The more reliable test we prepare the more effective in our testing system. As Heaton states, "reliability is a necessary characteristic of any good test: for it to be valid at all, a test must first be reliable as a measuring instrument" (Heaton 1991:162). Weir describes the concept of reliability as a fundamental criterion against which any language test has to be judged (Anastasi in Weir 1990: 31). He explains that the concern is "how far can we depend on the results that a test produces or, in other words, could the results be produced consistently" (Weir 1990:31).

Three aspects of reliability are usually taken into account. As Weir presents, the first aspect of reliability concerns the consistency of scoring among different markers. The second aspect refers to how to enhance the agreement between markers by establishing and maintaining adherence to, explicit guidelines for the conduct of marking. The third aspect of reliability is "of parallel-forms reliability, the requirements of which have to be borne in mind when future alternative forms of a test have to be devised" (Weir 1990:32).

Hughes suggests ways how to make tests more reliable. They are: take enough samples of behavior, do not allow candidates too much freedom, write unambiguous items, provide clear and explicit instructions, ensure that tests are well laid out and perfectly legible, candidates should be familiar with format and testing techniques, provide uniform and non-distracting conditions of administration, use items that permit scoring which is as objective as possible, make comparisons between candidates as direct as possible, provide a detailed scoring key, train scorers, agree acceptable responses and appropriate scores at outset of scoring, and identify candidates by number, not name, employ multiple, independent scoring(Hughes 2002:36-42)

So, even though we did not really touch the reliability of the test we tried to see the reliability of this exam in terms of some of the criteria put by Hughes (2002) as follow:

- Do not allow candidates too much freedom: in terms of freedom this exam is good because it did not give that much freedom rather it restrict students to answer the questions.
- Write unambiguous items and provides provide clear and explicit instructions: most of the items are clear and prices in this exam so this makes the reliable in terms of this sight.

- In terms of tests are well laid out and perfectly legible: the text as well as the questions are legible enough and the candidates should be familiar with format and testing techniques. The provide uniform and non-distracting conditions of administration and use items that permit scoring which is as objective as possible because all the exam questions are multiple choice
- Finally, agree acceptable responses and appropriate scores at outset of scoring, in this exam has commonly agreed answers.

So, based on the above reliability criteria designed by Hughes this exam is prepared in very well manner but in terms of it variety in item type and skill is not. And also in terms of scoring since we are not check the result of the students we did not recognize it perfectly.

Validity of the Test

Validity is one criteria of good test. As Heaton states, “the validity of a test is the extent to which it measures what it is supposed to measure and nothing else” (Heaton 1991:159). In other words, a test is said to be valid if it measures accurately what it is intended to measure. So in terms of these criteria we tried to see the validity of this test as follows:

In terms of content validity the exam is not valid in terms of content because the exam loss some of the contents in the book or the contents which need to be measure are not included in the exam fore exam in the book listening, writing activities, suffixes and prefixes, in grammar part all the contents are not covered so due to this problems this exam is not that much valid in content validity. Supporting our idea, Hughes claims “a test is said to have content validity if its content constitutes a representative sample of the language skills, structures, etc. with which is meant to be concerned” (Hughes 2002:22). So, the samples are not really representative in this exam.

Hughes presents two kinds of criterion-related validity: concurrent and predictive. Concurrent validity is when the test scores are correlated with another measure of performance; usually an older established test, taken at the same time (Weir 1990:27). Predictive validity concerns the degree to which a test can predict a candidates’ future performance. “If a test has construct validity, it is capable of measuring certain specific characteristics in accordance with a theory of language behavior and learning” (Heaton 1991:161). In terms of this criteria we cannot see the exam because we cannot check it in terms of the students test score but In Hughes’ words, the test has construct validity if it measures just the ability which it is supposed to measure (Hughes 2002:26) in this case the test try to be constructive since it measures different ability of the students even though there are different problems.

Hughes claims “a test is said to have face validity if it looks as if it measures what it is supposed to measure (Hughes 2002:27). He gives an example of a test that pretends to measure pronunciation ability but does not require the test to speak. As far as we concerned, in face validity we observed some problems like in : grammar part it covers some part of the grammar aspect like questions, tens, conjunctions, articles, prepositions, but in did not test other aspect of the grammar(prefixes/suffixes, conditional tens, voice etc) but these all contents are existed in the book. So, if the exam try to test all measurable contents this exam will be more effective.

Biasness of the Test

The exam we prepare to test students need to be free from biased in sex, religion, genetics etc. it must be see the students equal: so, we try to see this exam in terms of this criteria. In this exam paper there are five instructions: all the instructions have no biasness in every aspect. Whole in the question there is no any biasness that we observed.

Evaluation of the Instructions

In this exam there are five instructions so we tried to evaluate all the instructions in detail. Part one reading part says “read the passage. Then answer the questions below.” In this instruction we have seen different limitation first the word below indicates all the question in the exam so it is better to be from question no this to this and second the instruction should put in clear way in which it indicates the question is multiple answer question. Part two vocabulary a section called completion so the instruction is not doesn’t grammatically correct. And it may not be a good instruction as multiple-choice instruction. And substitution part in order to make clear to the students it is better if it says question number this to this rather than saying it the following sentences. Part three word order the instruction says “the following words are not in the proper order. From the given alternatives choose the best one that is corrected.” So this is not grammatically faulty sentences and the instruction doesn’t fully transfer the message. Part four grammar indicates the place where did the students put their answer so this instruction is better than others we seen before. Part five communicative “choose the most appropriate expression to complete the dialogue and write on the letter of your choice in the answer sheet” so this indicate only expression but there is a question which needs response so it is better includes bot. Especially question no 40 is differ in nature there is no dialogue.

In short even though the instructions are clearly put it should be better if there is clear space in between them, write in bold and indicate for each he number of question in which it belongs.

Quality of the Choices/ Distracters

In preparation of a test the quality of choices need to be qualified and effective enough: the choices need to be similar in size, length and they will make the students thinking critically. So, we also see the exam for multiple choices questions. All the choices are similar in every aspect so on this side the exam is good about the in part one reading comprehensions incorporates ten question this questions incorporate best destructors let us take one example:

Question 9: The underlined word “They” in paragraph three refers to

A. Benefits B. real life situation C. offers D. customers

So, when we see this question makes the students to answer them all the choice and think critically and they cannot easily guesses the answer.

Question 6: The word underway in paragraph 6 is closet in meaning to:

A. Permanent B. complete C. beneficial D. in progress

When we see this question all the destructor are not similar choice ‘D’ is differ from other so this may show a clue to the students so it should be better if it is similar.

While in part two also the vocabulary part the choices are best but we observe some problems. Question no 14 has four choices A, B, and D choices are Regular verbs while choice C is irregular verb but the choice needed to

be similar in each feature. Question no 12 also choice C. pleasure is noun while other choices are noun so the choice need to be similar.

Part three word order part is good they make the students do the question critically. In similar way grammar part is also good but in question '26' says

“The cloud is very heavy. It to rain, A. will B. may C. has D. is going

So, this question is good but the choices are good distracter (are not similar) so if the destructors are similar with the answer the students will be think critically and not able to get the answer easily.

Question 17: “have you seen his hand writing? It was so illegible every one wondered what it meant A. unreadable B. untidy C. visible D. Illegal

So, when we look this question even though it was good the destructors are not that much good destructors choice and there is a mixture of negative and positive for example the word visible is totally different from other choice so this gives the students a chance to leave choice automatically.

To generalize this part, in terms of the quality of destructors we did not see any big problem because to check the quality of destructors we need to make item analysis but as far as we see the general quality in terms of their similarity and difference in between is good.

Grading or Marking

In preparing a test we need to put the load for all the exams and for each instruction in terms of their difficulty. Because the students need to be aware of the load of their exams and they will predict their result. But, in this exam there is no any mark indicators. We observed this thing as the weakness of the exam.

Contextually and Authenticity of the Exam

In preparing language test the test should be prepared in context to mean that if the exams are prepared by considering students background (culture, social structure and geographical location) it will not be difficult to do. So when we see this exam, the exam is not that much isolated from the students' background but in our view the passage will be better if it is highly related to Ethiopian culture. While in question we observe some conceptuality. For example, question 27 says “Dr. AbiyAhemed this issues is highly related to students' context, while in question 23 the issues is African cup of nation but the intention is to as the propositions but the student will motivate to this question since they are aware of such things etc. and when we see the reading text, it is good the students may be familiar with the issue when they learnt other subjects like geography and physics so it is authentic.

Limitedness

The exam question should be short enough they could not be that much long. It should be short, precise and simple. But to mean short means the information wants to deliver on the question not to be stopped but the language would be short. So, the exam we select for evaluation is good in such manner. The questions are presented in good way as well as the choices.

Test Inter-Activeness

The other principles is need to concern testing issue is inter-activeness.

Inter-activeness refers to the extent and type of involvement of the test taker's individual characteristics in accomplishing a test task (Bachman & Palmer, 1996, p. 25). Due to the different areas of language knowledge, planning strategy and personality, how could we give each testee a fair chance is always a question. As far as we concerned in the exam this exam is not that much even though the exam is divided in to parts and each part has its own clear instruction the exam need to be put in more interactive way like the instruction of should be put in bold and it could be clear if there is a pace in between them.

Editing

Before the exam will be delivered to the examinee the question must be edited in correct way. Because it will help to reduce time we lost in correcting them at exam time, the student will be free from frustration in exam mistakes. So, in this exam we did not observed any problems which need to be correct so in terms of editing the exam it is good.

Economy of the Exam

Economy means a test is economical if its overall costs are manageable for the program. Those costs could be money but also include time involved, necessary equipment, and personnel. In this exam we tried to evaluate the economy of the exam in two way one in time usage for doing the exam and the wording of questions. As we see the exam in terms of time the exam has given sufficient time and in wording case this exam is also good we did not see any big problem.

CONCLUSIONS

Generally, this paper intended to address the evaluation of DebreMarkos Preparatory school grade 11th English second semester final exam. So, in depth evaluation by using different criteria's was done in doing so based on the results obtained from the following conclusions were drawn:

- As language test the exam should include/measure all skills.
- In terms of the item types the exam need to include all items in order to measure different intelligences of the pupils'.
- As per preparatory level exam, it should be better if it is nearly prepared in entrance exam level.
- In terms exam difficulty, the exam is as good as we observed.
- As per preparatory level the exam need to be measure all the level of intelligences of students.
- The questions also have clear, less biased and precise words.
- The exam did not that much covers the contents of the book but curriculum designer include the content for the sake of students language proficiency so even though it is impossible to includes all the contents basic and important things need to be include as a final exam.
- Since the language teaching approach shift from grammar translation method to communicative approach, the current method is CLT so the exam should be better if it is tens to more communicative rather than tending to grammar part.

In short, even though the exam tries to cover the content of the book and include variety of items and skills we have seen different limitations so, it is better if the exam prepared as grade 11th student exam.

REFERENCES

1. *Bachman, L and Palmer, A. (1996). Language Testing in Practice. Oxford: Oxford University Press.*
2. *Heaton, J.B. (1988). Writing English Language Tests (New Ed.). London: Dover*
3. *Hughes, A. (2003). Language Testing For Teachers (2nd Ed.). Cambridge, U.K: Cambridge University Press.*
4. *Madsen, H.S. 1983: Techniques in Testing. New York and Oxford: Oxford University Press.*
<https://doi.org/10.1177/026553228500200109>
5. *Weir, C.J. (1990). Communicative Language Testing. London: Prentice Hall International, Quoted in Mangubhai, F. (2004). Language Testing. Queen's land, Australia: USQ*