

THE INTERNATIONAL JOURNAL OF HUMANITIES & SOCIAL STUDIES

Investigating the Effectiveness of Reading Passages and Comprehension Questions in Grade 9 Teacher- Made Exams: The Case of ASCO Secondary School, Addis Ababa, Ethiopia

Ermias Mulatu

Lecturer, Department of English Language and Literature
College of Social Sciences and Humanities, Dire Dawa University, Dire Dawa, Ethiopia

Abstract:

This research was carried out to investigate the effectiveness of reading passages and comprehension questions in grade nine teacher-made exams of Asco Secondary School, Addis Ababa, Ethiopia. The data of the study were collected through document analysis, questionnaire and semi-structure interview. The data sources were three grade nine teacher- made reading exams of different years, ninety-six students and six English language teachers. The collected data were analyzed both qualitatively and quantitatively. The findings of the study revealed that the reading passages lacked coherence. Besides, the comprehension questions didn't focus on checking students understanding of the passages. Rather, they focused on certain areas like, looking for specific information and vocabulary questions. Besides, there were questions which can be answered without reading the passages. Moreover, it was found that most of the items could not discriminate. As a result, the majority of the items had unacceptable discrimination index. The conclusion made from the study was that the reading exams lacked some of the important constituents that make reading exams effective. Thus, it was recommended that giving teachers training or seminar on testing reading skills specifically and testing in general was significant for tackling the problem.

Keywords: *Effectiveness, requirements and techniques*

1. Introduction

1.1. Background of the Study

Testing according to McNamara (2000) is universal aspect of social life. Throughout history people have been put to a test to prove their capabilities. Tests to see how a person performs in a particular area have become very important.

When we come to the teaching learning process, "Both testing and teaching are so closely interrelated that it is virtually impossible to work in either field without being constantly concerned with the other" (Heaton, 1988:5). The effectiveness of teaching should be checked. The instrument used to do this is testing. So whenever we teach, testing enables us to know how successful our teaching is and how much we fulfill the objective.

Specifically, language test is a tool that tries to evaluate how much has been learned in a language course or some part of a course (Oller 1979). As to McNamara (2000) what is true of testing in broad is also true of language testing. However, language tests are generally very important for three main reasons. First, they can serve as a gateway in people's life by helping them in education, employment and while moving from one country to another. Secondly, language tests can support people in their professional life. For instance, language tests can serve administrators and teachers by being the base in their decision making, such as, the placement of students on particular course. Finally, while conducting research in language study, researchers may need to choose an appropriate existing language test or design their own, to measure the language proficiency of their subjects. In general, understanding the use of language testing is very vital for both test creators and users.

Teacher made tests are very significant because no one knows the students and the language areas which need focus in the testing as teachers does. Testing language skills is one part of language testing. Reading is among the four basic language skills. If the any kind of test including reading tests is not well prepared it will definitely be difficult to assess the learners' real performance. Therefore, tests themselves should be evaluated in order to check their effectiveness. Otherwise they can be the cause for ruining the measurement of students reading skills.

1.2. Statement of the Problem

If a teacher wants to measure the students reading skills effectively he/she has to set reading exam which fulfills the criteria set by scholars in the area. Thus, the reading passages and the comprehension questions should be prepared proficiently. However, when we come to our country's case, there are various problems in teacher-made reading exams. The problems must be evaluated in order to help teachers to improve it.

As a teacher, the researcher came across criticisms of students, teachers, school managements and educated parents in teacher made reading passages and comprehension questions. In general, the criticisms focused on the contents and structures of the texts and the questions. Besides while taking Language Testing course in Addis Ababa University, the researcher got the chance to notice reading exams full of problems. Some of the problems were: spelling errors, arrangement problems, length of passages and ambiguous questions. These rationales provoked the researcher to conduct this study.

Some local studies have been conducted related to reading skill. Mekonnen (2005) carried out a research and examined the cognitive levels of the reading questions and teachers' implementation. He concluded that the cognitive level of the reading questions in this text book are more of the lower level type. On the other hand, teachers are found to be aware of the case and making efforts to supplement the questions during the lesson.

Samuel (2012) conducted a research on factors that affect students' reading comprehension in two government high schools, by focusing on grade nine. He found out that students have reading comprehension problem due to bad home condition and poor reading habit. Using only loud reading and lack of preparation of teachers are also other factors affecting students reading comprehension. In a similar case, Sileshi (2007) investigated reading comprehension abilities of year 12 preparatory school and revealed that the students have low reading ability due to the time allotted for reading in earlier grades. In addition to this, he mentioned that teachers didn't update themselves to have recent knowledge of teaching reading skill. Taye (1999) in his dissertation examined the effect of academic background on reading comprehension. His finding mainly confirmed that students can be disadvantaged by being tested on texts outside their academic area. On the other hand, Gebretsadik (1997) investigated problems in the effective use of questions in teaching reading skills. He discovered that the reading passages in the text book are not guided by the three phases of reading questions (pre, while and post reading questions) and also teachers are dependent only on the questions found in the text book.

Some other researches such as, (Shambel, 2000) and (Nigussie, 2000) were conducted in the area of language testing. Shambel (2000) examined the effectiveness of final examination papers in relation to the new English text book of grade 11. He determined that various language skills are not included in the exams. Nigussie (2000) assessed the techniques of testing grammar and vocabulary. He revealed that very few techniques were used in the tests.

This research is different from the above studies because it focuses on evaluating teacher-made reading exams on the basis of the organization of the reading passages, item construction techniques used to construct the comprehension questions, effectiveness of the questions regarding assessing the students' reading skills, the difficulty and discrimination level of each item in the teacher made exams. As far as my reading is concerned, I couldn't find any research conducted in this area.

Even if it is difficult or impracticable to prepare perfect exams, yet being conscious of the common problems can help teachers to minimize these problems or at least reflect on them while evaluating the results.

The following research questions are set in an attempt to accomplish the objectives.

- How is the organization of the reading passages?
- Do the comprehension questions include various types of reading questions in order to assess students reading skills effectively?
- Are the reading passages and the comprehension questions up to the requirement?
- Do the reading exams incorporate different kinds of item construction techniques to compose the comprehension questions?
- How is the appropriateness of the distractors, level of difficulty and discrimination power of each item?

1.3. Objectives of the Study

1.3.1. General Objective

The general objective of this study is to investigate the effectiveness of English language reading passages and comprehension questions in grade 9 teacher made exams.

1.3.2. Specific Objectives

This study is specifically conducted to:

- evaluate the organization of the reading passages.
- determine whether the comprehension questions include various types of reading questions in order to assess students reading skills effectively.
- examine whether the reading passages and comprehension questions are up to the requirement.
- identify whether different kinds of item construction techniques are used to compose the comprehension questions.
- analyze the distractors, level of difficulty and discrimination power of each item.

2. Conceptual Frame Work

2.1. Testing Reading

Reading is not a single skill. It involves a variety of sub skills. As a result, English language teachers should recognize and include these sub skills with accordance to their purpose. Grellet (1990:4-5) names sub skills of reading that can be included in comprehension questions. These are: Recognizing the script of a language- Deducing the meaning -Understanding information when not explicitly stated-Understanding conceptual meaning-Understanding relations between the parts of a text through lexical and grammatical cohesion devices-Interpreting text by going outside it. -Distinguishing the main idea from supporting details-Extracting main points to summarize (the text, an idea etc.)

2.2. Guidelines for Producing Reading Comprehension Questions

Concerning preparing reading comprehension questions, Grellet (1990:6-8) and Hughes (1889: 154- 155) set some practical advice and considerations to be borne in mind when producing comprehension questions from reading passages.

1. Reading comprehension should not be separated from the other skills.
2. Reading is an active skill. It should involve guessing, predicting, checking, inferring, referring and asking oneself questions. Thus, this should be taken under consideration while preparing reading comprehension questions.
3. Variety of questions types should be used.
4. Be prepared to make minor changes to the text to improve an item.
5. It is important to use authentic texts.
6. Avoid simply matching a string of words in a question with the same string in the text.

2.3. Techniques of Reading Item Construction

There are different techniques for constructing reading exam items. However, if great care is not taken the methods or techniques used can cause problem in the students. Concerning this notion, Hughes (1989) indicates students may read perfectly but the techniques used to construct the questions can hinder them from answering the questions. Writing different item techniques effectively can help teachers to see learners reading skills efficiently and prevent the questions from hindering the students.

2.4. Item Analysis

Baker (1991) defines item analysis as a way of checking objective test items which contain problems or ambiguities in their construction and which escaped noticed during the construction of the test. Hughes (1989) also explains that even if individual items make their own contribution to the total test, some contribute more than others, and it is the purpose of item analysis to identify those that need to be changed or replaced. Similarly, Madsen (1983:180) explains item analysis as follow:

- Selection of appropriate language items is not enough by itself to ensure a good test. Each question needs to function properly; otherwise, it can weaken the exam. Fortunately, there are some rather simple statistical ways of checking individual items. This procedure is called "item analysis." It is most often used with multiple-choice questions. An item analysis tells us basically three things: how difficult each item is, whether or not the question "discriminates" or tells the difference between high and low students; and which distractors are working as they should.

The above idea clarify clarifies that preparing effective exam is not only about choosing appropriate language items but it also about preparing items which function properly.

Heaton (1988) states the three main uses of item analysis. These are: providing valuable information concerning: the performance of the students as a group, the performance of individual students; and the performance of each of the items comprising the test.

There are techniques used for investigating the performance of individual questions or items. According to Baker (1991) and Alderson et al (1995) facility value (F.V) and discrimination index (D.I) are the two most widely used techniques to calculate each objective test items. Facility value shows how difficult or easy a particular test item is and discrimination index measures the extent to which the results of individual item correlate results from the whole test.

2.5. Distractor Evaluation

Test analysis is not only about evaluating the stem but it is also about examining distractors. A distractor is an incorrect option, the job of which is "to distract the majority of poor students" (i.e. those who do not know the correct answer) from the correct option". (Heaton 1988)

Hughes (1989:162) clarifies evaluation of distractors in the following way:

- Where multiple choice items are used, in addition to calculating item-test correlations and facility values, it is necessary to analyze the performance of distractors. Distractors which do not work, i.e are chosen by very few candidates; make no contribution to test reliability. Such distractors have to be replaced by better ones, or the item has to be otherwise modified or dropped.

Moreover, Brown (2003:60) the efficiency of distractors is the extent to which (a) the distractors attract a sufficient number of test-takers, especially lower- ability ones (b) those responses are somewhat evenly distributed across all distractors.

According to Madsen (1983:184) "Weak distractors, often cause test questions to have poor discrimination or an undesirable level of difficulty. No set percentage of responses has been agreed up on, but examiners usually feel uneasy about a distractor that isn't chosen at least 1 or 2 examinees in a sample of 20 test papers". Ibid (1983) has identified the three common causes for this:

- Including an item that was drilled heavily in class – an item that almost everyone has mastered. Therefore, the answer is obvious; the distractors cannot “distract”.
- Sometimes a well- recognized pair is used (e.g. this/these, is/are, etc.). Even though not everyone has control of these yet, students know that one of the two is the right answer; no other choice seems likely. Here we need to choose another test format.
- A third cause is the use of obviously impossible distractors: (Did he do the work?/*A. Yes, he did. B.Birds eat worms. C.Trains can't fly.).

To conclude, item analysis is a means of checking each question of objective test items to make sure whether it functions properly or not. Facility value (F.V) and discrimination index (D.I) and distractors evaluation are the three common techniques to calculate each objective test items.

2.6. Criteria for Analyzing and Evaluating Reading Test Items

Hughes (1989) cited in Atkins et al (1996:154) states the following criteria for evaluating reading test items.

- Is there high probability of a student arriving at the correct answer by mere chance or guessing, or only a low probability?
- Does the item involve simply matching the question to the same word in the text?
- Will the wording of the item cause students to misunderstand it?
- Is the language used grammatically correct?
- Can the item be answered correctly by simply using general knowledge or common sense or is it necessary to demonstrate genuine comprehension in order to answer the question?
- Does the item involve words that are probably unknown to the students?
- Is the test item sufficiently contextualized?
- Do the questions in general check both literal comprehension and reading between lines?
- Overall is there a variety of question types in the test?

Alderson et al (1995) and Nuttal (1982) also mention criteria for analyzing reading exams.

- Are all the testes well edited?
- Are the questions varied in type?
- Do the questions test reading sub-skills effectively?
- Are there enough questions?
- Is the language of the item easier than the language of the text? If not, you are also testing understanding of the items.
- Do the answers require language that is too difficult for the students to handle?
- Is the item likely to be biased against or in favor of students of a particular genre /culture/ , back ground knowledge /interest?
- Are there several questions on every part of the text?

The above criteria set by Hughes (1989), Alderson et al (1995) and Nuttal (1982) were used as foundations for evaluating the effectiveness of reading passages and their comprehension questions.

To conclude, different scholars work related to the study were discussed in this chapter. Some of the issues raised here were: definition of reading, definition of language tests, testing reading, types of reading comprehension questions, guidelines for producing reading comprehension questions, techniques of reading item construction and parts item analysis. Finally, criteria for analyzing and evaluating reading test items were presented.

3. Research Design and Methodology

This study employed descriptive case study and both qualitative and quantitative methods. Qualitative approach was chosen to discuss the data collected from the interview with the teachers. Quantitative method was selected in order to analyze the questionnaire. Moreover, both quantitative and qualitative methods were used for the document analysis. The quantitative method used for the investigation of the exams and the item analysis (difficulty level and discrimination index) were numerically presents. Besides, the qualitative approach for the document analysis described and discussed the issue in words.

10% of the total populations were selected to make the sample manageable size. Therefore, from the total of nine hundred sixty-four students ninety-six of them were selected as a sample. Out of the 11 teachers grade 9 English teachers of Asco Secondary School, 6 of those who participated in the exam preparation were purposefully selected for the interview

The instruments of this study were document analysis, (the reading exam papers of the sample students) questionnaire and semi-structured interview

4. Results

4.1. Back Ground Information of the Reading Exams

As mentioned in the chapter three (3.3.1) this study focused on analyzing three reading exams from two different years (first and second and second semester exams of the year 2005 and first semester exam of the year 2006). The reading passages in each exam included different issues, number of paragraphs and words.

Exam A (first semester exam of 2005) contained four paragraphs and two hundred eighty-three words. It talked about development of countries and their people. It explained that development is reaching an acceptable standard of living and fulfilling the basic needs in which the people require. It also stated that developed countries are not necessarily rich and rich countries are not necessarily developed. It concluded by stating that development is the cooperative effort of different countries, institutions and it is an ongoing activity.

Exam B (second semester exam of 2005) had five paragraphs and two hundred seventy-seven words. It discussed about the Hunza people who live in a remote region of the high mountains of North Pakistan. Most of the Hunza people live over hundred years. They are among the healthiest people in the world. Their diet is so healthy. Grain, fruit and vegetables are their daily foods. They never drink alcohol. Sugar and canned food are unknown to them. Having short cooking time and not exposed to foods that causes overweight are the other reasons for their healthiness.

Exam C (first semester exam of 2006) contained three paragraphs and two hundred sixteen words and it narrates about a poor man and his rich neighbors. The cheerful poor man spent his time by singing. This routine made the rich ones unhappy. As result, one of the wealthy men threw a big bag of money to the poor man's house so as to stop him from being joyful. The poor man started to worry a lot. However, when the rich man told him that he intentionally threw, he became angry and gave back the money to the rich man. He finally gained his joy.

Each of the exams had ten comprehension questions. Thus, this study focused on examining a total of thirty reading comprehension questions.

4.2. Examining the Types of Reading Comprehension Questions Used

The following Table shows the proportions of the different types of reading comprehension questions observed in the three teacher-made grade 9 achievement exam papers of Asco Secondary School.

Question Types	Exam A	Exam B	Exam C	Total	Percentage
Literal comprehension	6	5	5	16	53.33 %
Reorganization or Reinterpretation	1	-	-	1	3.33 %
Questions of inference	1	-	2	3	10%
Questions of reference	1	1	-	2	6.67%
Questions of evaluation	-	-	-	-	-
Questions of personal response	-	-	-	-	-
Vocabulary questions	1	4	3	8	26.67 %
Total	10	10	10	30	100 %

Table 1: Types Comprehension Question Used and Their Proportion in the Exams

Note: exam A refers to the 2005 E.C first semester final exam, exam B represents the 2005 E.C second semester final exam and exam C stands for the 2006 E.C first semester final exam.

The exams dominantly used literal comprehension questions. Thus, they lacked variation of question in type which would help to check various reading skills of the students. The major aim of preparing reading passages and then questions out of the passages is in order to check candidates understanding of the text. However, the two dominantly used question types (literal comprehension and vocabulary questions) in their nature require candidates to search specific information than checking their understanding of reading passages.

4.3. Using Reading Exams Which Fulfill the Requirement

In preparing reading exams (passages and comprehension questions) teachers are advised to follow the criteria set by different scholars in the area. This research used the guidelines set by Grellet (1990:6-8) and Hughes (1889: 154- 155). (See chapter two: 2.9) and also Alderson et al (1995) Nuttal (1982) and Hughes (1989), (See chapter two: 2.12) in order to evaluate the passages and the comprehension questions.

4.3.1. Use of Simply Matching a String of Words

Reading exams are anticipated to test students reading ability. As Grellet (1990) clarifies, reading is an active skill which needs understanding of a text and it constantly involves guessing, predicting, checking and asking oneself a question. Yet, if the questions required simply matching the question to the same word or phrase in the passage, then students could answer the question without understanding the text. Hughes (1889) guides teachers to avoid preparing questions which require simply matching a string of words or phrases in a question with the same string in the text. However, the sample exams had questions which could be answered without understanding the passage.

- Exam B, No.3. They _____ eat butter and eggs.
A. Seldom B. continually C. often D. always

The stem was directly copied from the passage B paragraph 3 line 3 just by taking out the word 'seldom'. Any student who can read English could answer this question because it didn't require candidates' understanding of the passage. It could easily be answered by just matching the sentence in the question with the sentence in paragraph 3 line 3 of passage and fill the missing word which was 'seldom'. Besides, the stem itself was not that clear because the pronoun 'They' was contextualized only in the passage but not in the

question. Thus, it didn't clearly show what it refers to. The above question could be improved in the following way so as to make it capable of checking the candidates' understanding of the passage:

- How often do the Hunza people eat eggs and butter?

Moreover, in the alternatives it was better if the word 'seldom' was changed to its synonym 'rarely' in order not to give students a chance of matching the word in the passage with the same word in the alternatives.

There were also simply matching questions which directly copy longer sentences from the passage and require students to choose the correct answer just by taking out a phrase from it. For instance:

- Exam A.No.4. Developed countries are not necessarily rich and rich countries are.....
 - A. necessarily developed
 - B. not necessarily rich
 - C. not necessarily developed
 - D. all are correct answer

In this case, a phrase was taken out to be set as an alternative not a word. The stem (question) could directly be found in passage A paragraph 2 line 1.

Some of the simply matching questions were so direct that they vividly indicate the answer to the respondents in two ways: firstly by making it bold and secondly by forwarding direct meaning of the word. The following question can be a good example for this.

- Exam B.No.8. What does shangari_La mean in English?
 - A. paradise
 - B. remote
 - C. mountain
 - D. river

Passage B paragraph 1 line 1 and 2 said 'They called the place Shangari-La. This name came to mean "paradise" in the English language'. As it is clearly seen, the first sentence forwarded what the place is named and the second sentence presented its English meaning by making it bold and underlined. Thus, the answer is obviously 'paradise'.

Reading comprehension questions should be capable of checking students' understanding of reading passages. In other words, respondents should answer the questions after reading the texts and understanding them. In this regard, some of the items in the sample exams were prepared in a way that, they could not be answered without understanding the passages. For example:

- Exam C.No.5. Why was the poor man happy when he returned home?

- A) He got a full big bag of money
- B) He got a beautiful wife at his room
- C) That was his habit
- D) he did as he had done befem

- Exam C.No. 6. What made the poor man fear?
 - A) his poverty
 - B) he might be accused of having stolen money
 - C) his neighbours warned him to stop singing
 - D) loneliness

- Exam C.No.7. According to the writer's idea, what is the general idea of the passage?
 - A) being rich does not give happiness
 - B) Being poor makes people happy
 - C) Being both rich and happiness makes people happy
 - D) None

- Exam A.No.1. What talks paragraph one about?
 - A. Definition of development
 - B. The millennium development goal
 - C. The Cooperation of poor & developed countries
 - D. All are correct

Exam A.No.10. Write the possible title of the passage.....

- ✓ **Note:** all the questions were copied as they are from the original papers

Even if the above questions had their own problems (editing, grammatical, distractor choice, repetition of ideas) which will be discussed in the coming parts of the analysis, they were good regarding testing respondents' understanding of the passages. In other words, each of the above items checks the students' reading comprehension ability. For instance, students who didn't read and understand the passage would face difficulty in answering the cause for the poor man's happiness when he returned home? (Exam C.No.5), the reason that made the poor man fear (Exam C. No. 6) or they would not be able to write possible title for the passage (Exam A.No.10).

In summary, the sample exams contained some questions that required simply matching a string of words or phrases with the same strings in the passage. Yet, they also contained questions that could only be answered by reading and understanding the passages rather than by merely matching of strings of words or phrases.

4.3.2. Use of Reading Comprehension with Other Skills

It is hard to think of language skills separately. In a reading exam teachers can especially integrate two skills. These are reading and writing. Grellet (1990) explains that there are few cases in real life situation in which we carry out only one skill. We talk or write about what we have read. It is therefore important to bond the different skills through the reading exams. We can for example ask candidates to write the summary of the passage or to rearrange the jumbled sentences from a passage or a paragraph.

Regarding the sample tests there was only one question that incorporated reading and writing to some extent, i.e. Exam A.No.10 which asked candidates to write the possible title for the reading passage.

The reading passage in Test C which talked about the poor man and his rich neighbors was suitable for preparing comprehension questions which orders candidates to rearrange jumbled sentences from the passage and write them again. In this way the two skills could be integrated.

To sum up, integrating reading questions with other skills is vital. The sample reading exams had only one question that fulfills this criterion. Thus, it could be said that the sample exams are not integrated with the other skills.

4.3.3. Use of Grammatically Correct Passages and Comprehension Questions

In this section, every grammatical error in the sample tests had been closely examined. According to Hughes (1989) the language used in reading passages and comprehension questions should be grammatically correct. The incorrectness in grammar could make distractors less plausible (attractive). The sample exams had different grammatical problems. For instance:

- Exam B, No 4. There is a _____ area of farming land in the Hunza valley.

A. large B. limited C. Plentiful D. unknown

The article 'a' in the above stem gave a clue that option D couldn't be the answer because the article 'a' doesn't take a vowel sound word after it. Thus, due to the grammatical deviation between the stem and options, choice D has less chance to be selected as an answer. Concerning this, Heaton (1990) states that in designing multiple choice questions, if you use incomplete sentences it is important to make the entire stem and option grammatically correct. Grammatical mismatch between the stem and options would make distractors not to function properly.

In addition to the above example, exam A also has grammatical error that appeared in a stem with an incomplete sentence and one of its options.

- Exam A, No 3. According to the writer's view development is _____

A. limited task
B. endless task
C. happens only in developing countries
D. all

The stem ended with a phrase which said 'development is ____'. When one tried to complete the unfinished sentence with the alternative given, he/she could realize that option C didn't grammatically match with the stem. The complete sentence 'Development is happening only in developing countries.' is ungrammatical.

The reading passages also contained grammar errors. For example, passage A, paragraph 3 line 2-3 there is a subject verb agreement problem. The subject of the sentence is government but the verb used was try. Here is the example. 'The government of richest countries try to help ...' Since the word government is considered as third person singular noun, it should take the verb 'tries' not 'try'. Then it could be improved in the following way. 'The government of richest countries tries to help....' Or 'The governments of richest countries try to help...'

Exam C, No 1. option D 'he wanted to need some money', exam Thestem, 'What talks paragraph one about?' and exam A. No 8 the stem 'The World Bank provide money...' were some of the other grammatical errors observed in the passages and questions.

In general, grammatical errors in reading exams could make distractors not to work as they should and they could also be an obstacle for the pure flow of ideas in reading passages. In addition to this, they could give hint to the correct answer.

4.3.4. Using Well Edited Tests

Tests in their nature should be well edited (Alderson et al 1995). Reading tests (reading passages and comprehension questions) should also be well edited because something which is not edited in the passage can cause problem in the comprehension questions. There were different editing problems in the sample exams. Let us see the following examples.

Spelling errors were one of the editing problems seen in the sample exams. They could be obstacles in understanding the passages and the questions. For instance, exam C the title states 'Being rich does not give haypilles'. It was meant to say 'Being rich does not give happiness' in addition to this, exam C paragraph 1 line 2, the word 'annoys', exam A No3. in the stem part the word 'according', and exam C Paragraph 2 line 2 the word 'neighbors' were some of the other spelling errors.

The other editing problem was related to punctuation and capitalization. Exam A Paragraph 4 line 2 for example contained a sentence with wrong placement of punctuation and capitalization. 'People always try to improve the quality of their live. And the lives of their children.' There was no need to use full stop between the word 'live' and the conjunction 'and'. It could be written in the following way. 'People always try to improve the quality of their live and the lives of their children.' Moreover, exam B in the reading passage paragraph 3 line 2 had both punctuation and capitalization problem '.... The Hunza diet. Is so healthy.' The full stop between the word 'diet' and the word 'is' was not necessary. Besides, capitalizing the next letter 'I' was wrong. Thus, it could be improved as follow: '.... the Hunza diet is so healthy'.

Missing (not including) words, was the other category of editing problem in the tests. Test A, no 7 both alternative C and D needed the relative pronoun 'who' after the word 'people'. (See appendix A). There were sentences lacking prepositions. For instance, test A paragraph 3 line 1 'The government poor countries work to improve...' missed the preposition 'of' between the word 'government' and 'poor'.

Replacement of words with other wrong words was the last classification of editing problem in the sample exams such as exam C, no 7 option C. 'Being both rich and happiness makes people happy.' In this case, 'happiness' was the word in the wrong place. It was intended to say 'Being both rich and poor makes people happy'

As said by Alderson et al (1995) exams should be well edited. Any type of exam which is not well edited can create ambiguity. As a result, the editing issue shouldn't be considered as a minor matter because it could hinder respondents from answering the questions.

4.4. Some General drawbacks of the Reading Exams

Not mentioning sources of the reading passages was the other problem of the exams. Mentioning the sources could create a chance for students further reading on the document and it also shows that the exam writer(s) gave credit for the for the author(s) of the text. In exams, it is advisable to state how much mark (value) each item has. Nevertheless, none of the sample exams mentioned the mark in which the comprehension question contained.

The instruction of all the sample exams didn't offer enough idea for the students on what they are going to do and how they are going to do it. Are they going to underline or circle on the answer? Specifically, exam A, contained two types of questions which were multiple choice questions and fill in the blank question under one instruction. (See Appendix A).

Generally, the sources of the passages and the mark each item contained were not mentioned. In addition to this, using different kinds of questions under one instruction was observed.

4.5. Analysis of Item Construction Techniques Employed

Item Types	Exam A	Exam B	Exam C	Total
Multiple Choice	9	10	10	96.7%
True/False	-	-	-	-
Matching	-	-	-	-
Completion	1	-	-	3.3%
Cloze	-	-	-	-
Information Transfer	-	-	-	-
Ordering Task	-	-	-	-
Editing	-	-	-	-
Total	10	10	10	100%

Table 2: Distribution of Item Construction Techniques Employed

Even if multiple choice questions have their own advantages, Hughes (2003), cited in Brown (2003:55) points out a number of weaknesses of multiple-choice items. Such as, they test only recognition of knowledge, guessing may have a considerable effect on scores, they severely restrict what can be tested, it is very difficult to write successful items, wash back may be harmful and cheating may be facilitated. Thus, it is possible to conclude that mixing various item construction techniques in testing students reading ability is vital.

4.6. Item Analysis

4.6.1. Analysis of Difficulty and Discrimination Levels

Exam	Easy	Difficult	Moderate	Need to be rewritten	Total
A	8	--	1, 3, 4, 5, 6, 7, 10	2, 9	10
B	10	7	1, 3, 4, 6, 8, 9	2, 5	10
C	3	8, 10	2, 4, 5, 6, 7, 9	1	10
Total	10%	10%	77%	13%	100%

Table 3: Difficulty Level of the Three Exams

On the basis of the students' results, the items were classified as easy, difficult, moderate, and need to be rewritten. As table 3 above shows most items 77% of them were moderately difficult, while 10% of the items were easy. 10% of them were difficult. The rest 13% needed to be rewritten.

Exam	Not acceptable	Acceptable	Questionable	Total
A	2, 4, 5, 8, 7	3, 6, 9, 10	1	10
B	3, 6, 7, 10, 9	1, 4, 5, 8	2	10
C	2, 3, 8, 5	4, 6, 9	1, 7, 10	10
Total	47%	36%	17%	100%

Table 4: Discrimination Index of the Three Exams

As table for above pointed out, 47% of the items were unsuccessful in discriminating the students. Thus, they had unacceptable discrimination power. Meanwhile, 36% were successful in discriminating students' achievement. Therefore, they were categorized acceptable discrimination index. The rest 17% were questionable, which means they incorporated some form of errors (inaccuracy).

5.1. Summary of Major Findings and Conclusions

All the sample reading passages had unified paragraphs in which the detailed sentences were directly related to the topic sentences. Having unified sentences makes a paragraph clearer and easy to read. However, some of the passages had problem of coherence in which the sentences were not well linked. More than half 77% of the items were moderately difficult. Yet, there was a problem of selecting appropriate distractors. Regarding the discrimination index, majority of the items 47% had unacceptable discrimination power. Direct matching questions, inconsistent alternatives and response clues were the reasons for the unacceptable discrimination power. Although using different types of comprehension questions helps to assess the various reading skills of the students, majority 53% of the types of reading comprehensions employed were literal comprehension questions. The other types reading comprehension (see Table 1) which were very vital in testing students reading skills were given much less or no attention. In short, unfair distribution of reading question types was employed. Reading comprehension tests should fundamentally check respondents understanding of the text presented. However, there were some questions which could be answered without understanding the passage just by matching string of words or phrases in the passage. Besides, there were grammatical and editing problems in both the passages and comprehension questions. Moreover, none of the sample tests integrated reading with other skills and the sources of the passages were not mentioned in any of the exams. While there were various item construction techniques which could be used in order to check the students reading skills, the sample exam dominantly used multiple-choice items. Out of the total of thirty questions 29 of them 96.66% were multiple choice questions. Moreover, due to the nature of the technique used, some of the questions could be answered just by detecting the response clues given by the options. As the questionnaire to the students demonstrated, majority 80.21% of the respondents believed that the instructions of the passages were clear but more than half 63.54% of them stated that they were not interested in the sample passages. This revealed that the students were not interested in the passages but they appreciated the instructions. Moreover, 58.32% of the students reflected that there are questions which could be answered without the need of understanding the passage. As the semi-structured interview revealed 4 out of 6 of the teachers reported that they adapt reading passages from other sources. However, none of them mentioned the sources of the passages with a believe that it is not that important, being careless and due to coping the passages from other exams which didn't mention sources. This shows that the teacher tried to adapt passages without mentioning sources.

5.2. Recommendations

To alleviate some of the problems seen in the reading passages and comprehension questions of grade 9 teacher made exams, it is important to offer solutions. Thus, the following recommendations have been made based on the above drawn conclusions.

It is advisable to be cautious while selecting passages for reading exams. Thus, it is worthwhile to adapt and use reading passages which are coherent, interesting, well edited and grammatically correct. In addition to this, the passages should incorporate new ideas as much as possible.

It is wise for teachers to include various types of reading comprehension questions that require different reading skills.

It is suggested to teachers to use reading comprehension questions which could be answered only by understanding the passages rather than by directly matching a string of words or phrases and it is preferable to integrate reading questions with other skills.

It is desirable to use different item construction techniques rather than being dependent on a single kind of technique i.e. Multiple-choice.

It is advantageous for teachers to develop a practice of reviewing students' exams after the results are checked, so as to see whether the distractors work as they should or not. This will help the teachers for improving items in the next exams.

It is valuable to give teachers training or seminar specifically on testing reading skills or testing in general is significant for tackling the problem.

6. References

- i. Alderson, J.C. Clapham, C. and Wall, D. (1995). *Language Test Construction and Evaluation*. Cambridge: Cambridge University Press.
- ii. Atkins, J. Hailom Banteyirga and Nuru Mohammed. (1996). *Skills Development Methodology. Part 2*. Addis Ababa: Addis Ababa University.
- iii. Baker, D (1991). *Language Testing: A Critical Survey and Practical Guide*. (2nd ed) London: Athenaeu Press.
- iv. Brown, H. D.(2003). *Language Assessments Principles and Classroom Practices*. London: Longman.
- v. Brown, J. D. (1996). *Testing in Language Programs*. Upper Saddle River, NJ: Prentice Hall Reagents.
- vi. Gebretsadik Basen (1997). *Some Problems in Teaching Reading Skills in Grade Eleven: Unpublished BA Paper*. Addis Ababa: Addis Ababa University.
- vii. Grellet, F. (1990). *Developing Reading Skills: A practical Guide to Reading Comprehension Exercise*. (10th ed) Cambridge: Cambridge University.
- viii. Heaton, J. B. (1988). *Writing English Language Tests*. London: Longman.
- ix. Heaton, J. B. (1990). *Classroom Testing*. New York: Longman.
- x. Hughes, A. (1989). *Testing for Language Teachers*. New York: Cambridge University Press.
- xi. Madsen, H.S. (1983). *Techniques in Testing*. New York: Oxford University Press.
- xii. McNamara, T. (2000). *Language Testing*. New York: Oxford University Press.
- xiii. Mekonnen Gemechu. (2005). *The Cognitive Levels of Reading Questions and Teachers Class Room Supplement to the Questions*. Unpublished MA Thesis. Addis Ababa: Addis Ababa University.

- xiv. Nigussie Sinshaw (2000). 'The Testing of Grammar and Vocabulary in Grade 10'. Unpublished BA Paper. Addis Ababa: Addis Ababa University.
- xv. Nuttal, C. (1982). Teaching Reading Skills in a Foreign Language: Practical Language Teaching. London: Clays Ltd.
- xvi. Oller, J. (1979). Language Tests at School. New York: Longman.
- xvii. Samuel Moges ,(2011). Factors that Affect Students' Reading Comprehension in two Government High Schools. UnPublished MA Thesis. Addis Baba. Addis Ababa University.
- xviii. Shambel Tesfa ,(2000). Effectiveness of the Final Examination Papers in Relation to the New English Text Book Grade Eleven. Unpublished BA Paper. Addis Ababa: Addis Ababa University.
- xix. Sileshi Kassahun. (2007). An Investigation of Preparatory Students Reading Comprehension Abilities. Unpublished MA Thesis. Addis Ababa: Addis Ababa University.
- xx. Taye Regassa(1999)The Effect of Academic Background on Reading Comprehension. Unpublished Dissertation. Addis Ababa: Addis Ababa University.