



# **AMERICAN JOURNAL OF ARTS AND HUMAN SCIENCE (AJAHS)**

**ISSN: 2832-451X (ONLINE)**

**VOLUME 1 ISSUE 2 (2022)**

INDEXED IN



**PUBLISHED BY: E-PALLI, DELAWARE, USA**

## Test Construction Principles Violated in Selected Summative Tests in English

Jay-ar Cristobal<sup>1\*</sup>

### Article Information

**Received:** June 09, 2022

**Accepted:** June 20, 2022

**Published:** June 22, 2022

### Keywords

*Principle, Summative, Test, Violation*

### ABSTRACT

Teachers give many forms of assessment to measure their students' performance and achievement. Paper and pencil test might be considered as an old and traditional form of assessment, but it is still effective until now. However, it has been observed that many test papers contain erroneous test papers. Thus, this descriptive study aimed to determine the test construction principles violated in sixteen summative tests in English gathered from high school teachers in a public secondary school in San Nicolas, Ilocos Norte, Philippines. Frequency counts and percentages were used to treat the data gathered. Interpretation and analysis of data were guided by the test construction principles used in the study of Lin (2002). Results of the study show that most of the vocabulary items violated the principle on making all the options with the same word class. On the other hand, the majority of the grammar test items used mixed options. When it comes to the reading comprehension section, most of the items can be answered from general knowledge. These test construction violations, in many ways, could hamper effective and successful assessments.

### INTRODUCTION

Assessment has been a part of every educational system. Different kinds of assessments are given to students not only to measure their understanding about the topics and concepts of the lessons they need to learn but also to improve teachers' strategies and resources in the teaching-learning process. That is why teachers continue to look for effective ways in measuring their students' performance.

Paper and pencil test might be regarded as the most traditional way of assessing students' learning, however, it is still effective up to now. In language classes, teachers administer paper-and-pencil tests to assess separate components of language knowledge such as grammar, vocabulary, and reading comprehension. Considerably, these tests are commonly constructed through multiple-choice.

Linn & Gronlund (1995) contended that multiple-choice item is widely used by teachers, schools, and assessment organizations and is one of the most commonly applicable test techniques to assess achievement. The chief advantage of multiple-choice tests is that scoring can be simple, rapid, and economical.

Begum (2012), likewise, asserted that multiple-Choice questions have long been the preferred method and most commonly used in classrooms, providing for both reliable and valid coverage of content.

Although multiple-choice items test only recognition knowledge rather than the ability to use language, they are confirmed effective in assessing student's ability to recognize correct grammatical forms and vocabulary, comprehend a certain reading text, and make important discrimination in the target language (Bailey, 1998).

With these views, it can be said that multiple-choice has been widely used in language testing specifically in testing grammar, vocabulary and reading comprehension.

Nevertheless, a successful multiple-choice item is time-consuming and hard to construct. Considering its demands like time and effort, multiple-choice item is at times ignored and is not taken seriously by test developers which can only result to poor construction of tests.

While it is true that language teachers construct tests with multiple-choice items, the rules and principles in making such type of test are not followed. Thus, the poor quality of test papers may significantly affect students' scores in tests and ultimately their academic performance. Therefore, it is essential to follow certain rules and principles in constructing test items to measure what the teacher really intends to measure and to achieve his/her learning goals and objectives. These observations and views relative to poor test construction prompted the researcher to investigate the common errors in multiple-choice items in vocabulary, grammar and reading comprehension tests and to analyze its implications to language testing.

This study sought answers to the following questions:

1. What are the test construction principles violated in the selected summative tests in:
  - 1.1. vocabulary test;
  - 1.2. grammar test; and
  - 1.3. reading comprehension test?
2. What is the implication of the violated principles to test construction?

The conceptual framework as shown in Figure 1, served as the guide in conducting this study. Using the collected test papers from various sources, the violated principles in grammar, vocabulary and reading comprehension test items were identified and analyzed. The test construction principles violated in the test items served as the basis in generating the implications of such violations to test construction.

<sup>1</sup> Mariano Marcos State University, Philippines.

\* Corresponding author's e-mail: [jacristobal@mmsu.edu.ph](mailto:jacristobal@mmsu.edu.ph)

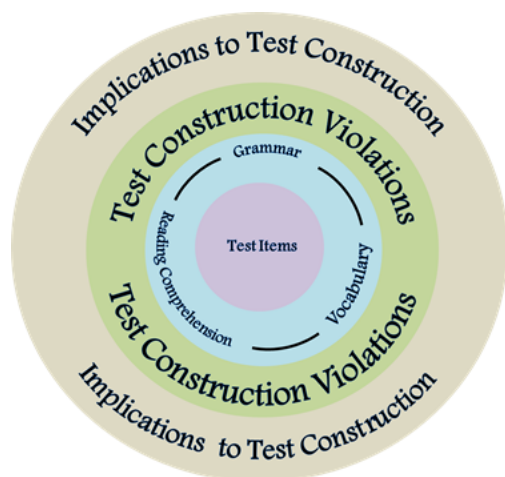


Figure 1: Research Paradigm

## LITERATURE REVIEW

### A Brief Background of Assessment

In the Philippines, the K to 12 Basic Education Curriculum (K to 12 BEC, 2010) aims to help learners acquire highly-developed literacy skills that enable them to understand that English language is the most widely used medium of communication in Trade and the Arts, Sciences, Mathematics, and in world economy. Furthermore, the curriculum aims to help learners understand the social conditions, and that English is inextricably involved with values, beliefs and ways of thinking about the world. Through multi-literacy skills, learners are able to appreciate and be sensitive to socio-cultural diversity and understand that meaning of any form of communication that depends on context, purpose, and audience. Relatedly, part of achieving these goals is improving and enhancing students' grammar, vocabulary and reading comprehension skills. One way of achieving such is through the use of effective assessments.

Brown (2004) presented five principles of language assessment such as validity, reliability, practicality, authenticity and washback. Aside from these, Bachman & Palmer (1996) likewise proposed six other important principles of language assessment which include reliability, construct validity, authenticity, interactiveness, impact, and practicality. These principles, if followed by test developers, could be of big help in attaining effective and successful assessments. Relatedly, Rubin, Daly, McCroskey, and Mead (1982, as cited in Sharma, 2020) suggested that well-managed assessment practice can have a desirable effect on teachers' instructional activities and test achievements can have a progressive function in students' learning.

Madziyire (2010) argued that teacher-made tests could help teachers to identify content knowledge or skills that have been mastered by students and that teachers know through the results of their tests in areas in which students experience difficulties. Consequently, they find ways on how their students could overcome such difficulties so that these students can do better. Likewise, the results of teacher-made tests allow teachers to assess both their

strengths and weaknesses. Therefore, well-constructed teacher-made tests could give students the opportunity of assessing their knowledge and through immediate and constructive feedbacks, the students can improve their performance (Sax, 2007).

In another perspective, Mpofu (2011) posited that while teacher-made tests help improve the performance of children in the learning process, it is not always the case since tests are constructed by teachers themselves. At times, these tests lack validity and reliability. It is clear that teachers need to be extremely careful in designing the test that measures the skill it intends to measure.

## MATERIALS AND METHODS

The research aimed to determine the common errors in multiple-choice items in vocabulary, grammar and reading comprehension tests taken from the items of the collected test papers. This study focused only on the errors with regard to the options in the test items. Sixteen summative test papers with 100 items each, were used to gather the data on the errors in multiple-choice items in vocabulary, grammar and reading comprehension. These test papers were constructed by high school English teachers in a public secondary school in San Nicolas, Ilocos Norte, Philippines. This number was considered adequate to represent the test papers with regard to errors in vocabulary, grammar and reading comprehension test items.

Table 1 shows the number of test items in vocabulary, grammar and reading comprehension sections of the test papers gathered. It can be deduced from the table that there are 85 reading comprehension items, 60 vocabulary items and 55 grammar items which served as the study's source of data.

Table 1: Distribution of collected test items in vocabulary, grammar and reading comprehension tests.

Sections	Number of Test Items
Vocabulary	60
Grammar	55
Reading Comprehension	85
Total	13

The data on test construction errors were gathered using the principles for writing multiple-choice items in vocabulary, grammar and reading comprehension tests used in the study of Lin (2002). The principles for writing multiple-choice items in vocabulary are 1) All options should be the same word class; 2) Grammatical clues should be avoided in the stem; 3) All options should be on nearly the same level of difficulty; 4) All options should avoid needless redundancy; 5) Avoid using distractors with similar meanings; 6) Avoid using distractors with opposite meanings; 7) Each item should contain one correct or obviously best answer; and 8) All options should be nearly the same length. On the other hand, the principles used in determining the errors in grammar tests are 1) Avoid using distractors that

are pronounced alike; 2) Avoid using mixed options; 3) Avoid nonsense distractors; and 4) Avoid using inconsistent distractors. Moreover, the principles for writing reading comprehension tests are 1) Items should not be answered by purely matching of words; 2) Items should not be answered from general knowledge; and 3) All options should be roughly the same length. Frequency and percentage distribution were used to treat the data. The errors obtained from the test papers were analyzed and interpreted. Finally, the implications of the errors identified to test construction were generated.

## RESULTS AND DISCUSSION

Test items should be student-friendly. It means that tests should not create fear among students. However, making them too easy by using options that violate the principles of test construction could only confuse and mislead the examinees. This part presents the analysis and interpretation of the errors in multiple-choice items in vocabulary, grammar and reading comprehension tests and the implication of these errors in language testing.

### Errors in Vocabulary, Grammar and Reading Comprehension Tests

This study determined the number errors in vocabulary, grammar and reading comprehension from the collected test papers. These are presented in Table 2A. It can be deduced from the table that 57 (67.05%) of the total items have errors in reading comprehension. This indicates that in constructing test items, the teachers made the highest number of errors or mistakes in reading comprehension as evidenced by its number and percentage of errors.

It can be noted from the same table that 20 (36.36%) of the total items in vocabulary have errors in that section while 15 (25%) have errors in vocabulary tests. The results indicate that the majority of the test items have errors in reading comprehension section. This can be attributed to the teachers' lack of knowledge in constructing good

**Table 2A:** Distribution of errors in vocabulary, grammar and reading comprehension tests.

Sections	Test Items (f)	Errors (f)	Percentage (%)
Vocabulary	60	15	25
Grammar	55	20	36.36
Reading Comprehension	85	57	67.05

test items. Although this type of test has been commonly given to students, there are still many teachers who are not properly guided with the principles in making such type of test. This finding contradicts the study of Lin (2002) which showed the highest number of errors in the vocabulary section.

### Test Construction Principles Violated in Vocabulary Test

Another concern of the study is to determine the principles violated in constructing vocabulary test as shown in Table 2B. It could be gleaned from Table 2B that 6 (40%) of the total errors violate the principle All options should be the same word class. This indicates that most of the items in vocabulary show errors on word class. Consider the following items that show this violation:

#### (V1) An establishment looms in Singapore every now and then.

A. appears                      B. grows                      C. recede

It can be observed in item V1 that among the three choices, the only word that does not agree with the subject is option C. The inclusion of this option could make students easily determine that this is not the correct answer, without any knowledge on the meaning of the underlined word, because the said option violates the rule on subject-verb agreement. Likewise, students might notice that option C is the only word which does not end

**Table 2B:** Distribution of violated principles in vocabulary test.

Principles	Test Items (f)	Percentage (%)
All options should be the same word class	6	40
Grammatical clues should be avoided in the stem	0	0
All options should be on nearly the same level of difficulty	0	0
All options should avoid needless redundancy	1	6.66
Avoid using distractors with similar meanings	2	13.33
Avoid using distractors with opposite meanings	2	13.33
Each item should contain one correct or obviously best answer	0	0
All options should be nearly the same length	5	33.33
Total	15	100

in -s, therefore, they may either drop or select that option. This makes the item erroneous since it is not consistent with the form of verbs in its options.

(V2) Lolita gives her predicaments on the issue to stop people from talking.

A. opinion of the people

B. side of the issue

C. conflict of the issue

D. proved the allegations

(V3) Out from the marsh, from the foot of misty hills and bogs, bearing God's hatred Grendel came.

A. soft wet ground



- B. desert  
C. fertile  
D. barren

Item V2 also violates that rule. It is interesting to note that in the said item, options A, B and C are noun phrases while option D is noticeably the only verb phrase. In this item, the students might perceive that a verb phrase is necessitated, so they might select option C without carefully examining what the phrase really means. This finding can be associated to the study of Lin (2002) which found a great number of items in vocabulary that violated the principle on giving the options with the same word class. It can also be deduced from the table that 5 (33.33%) of the total errors in vocabulary violate the principle All the options should be nearly the same length. Consider the item which violates the said principle.

In item V3, it is noticeable that the options are not of the same length. Examining the options of this item, there is a tendency that the longest option soft wet ground is the obvious answer because it gives a better detail about the underlined word than the other options. Thus, students might recognize easily that either of the options with the same length is not the right answer. On the other hand, students might disregard the longest option as well and just rely from the short ones as the right answer. To some test takers, they hold on to the notion that the longest option is the right answer, especially when they really do not know the correct one. However, it must be taken that the correct answer should not be generally the longest one.

Moreover, 2 (13.33%) have errors on using distractors with similar meanings and using distractors with opposite meanings respectively. This implies that the items contain options which are closely related or have similar meanings and options which have opposite meanings. The following items show options with opposite and similar meanings:

(V4) And then the lover, sighing like a furnace with a woeful ballad.

- A. cheerful  
B. sorrowful  
C. wonderful  
D. joyful

(V5) The house cost an arm and a leg to be constructed.

- A. very expensive  
B. very cheap  
C. hard to build  
D. hard to maintain

In item, V4, if students perceive that cheerful and joyful are close in meaning, they might eliminate both and simply select either sorrowful and wonderful as the correct answer. Considering item V5, students who recognize very expensive and very cheap as opposite words might select one of the antonyms of the underlined word as the correct answer, and eliminate the last two options hard to build and hard to maintain as the correct answer.

Considerably, it could be gleaned from the same table that there are no violations on the principles Grammatical clues should be avoided in the stem, All options should be on nearly the same level of difficulty and Each item should contain one correct or obviously best answer. This indicates that the teachers are aware and knowledgeable of these principles, thus, they strictly follow such rules in constructing vocabulary test items.

### Test Construction Principles Violated in Grammar Test

This study also determined the principles violated in grammar test as reflected in Table 2C. When it comes to grammar, 11 (55%) of the total errors are using mixed options. This implies that most of the grammar items violate this principle. Since grammar test is easier to construct, its quality is compromised. Examine the following item that shows such error:

**Table 2C:** Distribution of violated principles in grammar test. (n=20)

Principles	Test Items (f)	Percentage (%)
Avoid using distractors that are pronounced alike	0	0
Avoid using mixed options	11	55
Avoid nonsense distractors	7	35
Avoid using inconsistent distractors	2	10
Total	20	100

(G1) They (participates, will participate, had participated) in the nutri-jingle contest next week.

It can be seen from item G1 that it is used to test tense consistency. However, it is noticeable that the options confuse the notion of agreement with that of tense. Examinees might disregard participates automatically due to lack of subject-verb agreement, not on the basis of consistency of tense. In this case, the focus of the test item is not clear whether it really tests tense consistency or subject-verb agreement. As what Allanson and Notar (2019) posited, educators have an obligation to ensure that the assessments they create accurately measure

learning objectives.

Interestingly, 7 (35%) of the total errors in grammar are using nonsense distractors. Sometimes, teachers construct easy items and one way to achieve such is the use of nonsense distractors. The following items violate such principle:

(G2) I (missed, will miss, will missed) you when you leave for Canada next month.

(G3) Mr. Reyes (talked, will talk, took, talks) to you tomorrow

Considering the options in item G2, it can be noticed that the item contains structure that does not occur

in grammatically acceptable English. Will missed is obviously grammatically incorrect, so students can easily disregard it from the options. However, there is also a tendency that missed might be chosen since it is only one word compared to the two-word options will miss and will missed. This is an indication that using grammatically incorrect word/s can never be a good option as it only creates misconceptions about the rules of English grammar. Likewise, in item G3, took is considered a nonsense distractor because it is formed from another word. Talked, will talk, and talks are formed from verb talk, thus, students can immediately determine that took is not the correct answer. Just like items G1 and G2, item G3 also tests tense consistency. However, with the inclusion of took in the options, it can be disregarded not on the basis of tense but on diction or choice of

word. Therefore, nonsense distractors tend to be weak distractors and students might only learn errors from the test itself. In relation to this finding can be linked to what Begum (2012) and Hansen & Dexter (1997) claimed that good multiple-choice items depend on effective distractors and best assessed in terms of the quality of such.

### Test Construction Principles Violated in Reading Comprehension Test

Determining the principles violated in reading comprehension test is also a concern of this study as reflected in Table 2D. With regard to reading comprehension, 34 (59.64%) of the items can be answered from general knowledge. These items violate the principle Items should not be answered from general knowledge. Consider the following items that violate this principle: Mordred

**Table 2D:** Distribution of violated principles in reading comprehension test. (n=57)

Principles	Test Items (f)	Percentage (%)
Items should not be answered by purely matching of words	12	21.05
Items should not be answered from general knowledge	34	59.64
All options should be roughly the same length	11	19.29
Total	57	100

makes himself King of England and incestuously claims Guinevere as his wife. Guinevere escapes to the Tower of London. The Bishop of Canterbury reproaches Mordred for his usurpation and would-be incest, and when Mordred tries to kill him, he flees and becomes a hermit. Mordred wins many Englishmen to his side, then meets Arthur at Dover but is forced to retreat from him. In this battle Gawain is mortally wounded. As he dies he admits to Arthur that if it were not for his insane pride in insisting on unjust revenge, Launcelot would be here now to save the kingdom; then he writes Launcelot, begging him to come help Arthur and also to pray at his tomb. Then, bleeding from the wound he got originally from Launcelot — with the fated sword of Balyn — Gawain dies.

(RC1) Based on the selection, who made himself King of England?

- Gawain
- Mordred
- Launcelot
- Arthur

(RC2) Where did Mordred meet Arthur?

- Dover
- Lover
- Cover
- Sober

(RC3) According to the selection, who died in the battle?

- Gawain
- Mordred
- Launcelot
- Arthur

(RC4) According to Gawain, what is the reason why Launcelot was not there to save the kingdom?

- Gawain's insane pride

- Arthur's insane pride
- Launcelot's insane pride
- Guinevere's insane pride

It can be noted that the given items RC1, RC2, RC3 and RC4 test or measure general knowledge only. It can be observed that the questions are under the literal level which means that the answers can be found in the reading text, itself. Students can, therefore, choose the answer correctly without understanding the whole passage by just simply locating what the questions are asking for. The answer is general knowledge for most people, regardless of the content of the passage. Consequently, these items do not develop higher thinking skills because the answers can be immediately found in the passage. As what Henning (1982) posited, in reading comprehension, items should require understanding of the passage. It should not be feasible to answer items correctly purely on the basis of general knowledge.

It could also be gleaned from the table that 12 (21.05%) reading comprehension items violate the principle Items should not be answered by purely matching of words. This indicates that the correct answers in these items merely match the same words in the passages. The following items can be answered by purely matching of words:

What is faith? Faith is believing in possibilities. It is the ability to carry on with our plans or to be true to our work even when though we feel discouraged or tired. It is staying active in relationship even when our friends aren't able to respond.

If there were no doubts, there would be no need for faith. Faith is temporarily putting our doubts on the shelf and working toward our goals Faith is trusting that help and support will be there for us even though they are not

in view. It is like looking at a map and choosing a new destination, getting on the road to go there, and trusting that the marks on the map symbolize a real place that we will find.

(RC5) According to the author, what will happen if there is no doubt?

- a. There would be no need for faith
- b. There would be no relationships
- c. People would get discouraged
- d. Confusion would occur

(RC6) According to the selection, having faith is like

- a. opening a book
- b. looking at a map
- c. driving a car
- d. writing a diary entry

It can be noticed that items RC5 and RC6 do not develop comprehension because the answers are solely taken from the passage. Thus, the correct answer can be easily determined because the it is copied exactly from the reading text. With this, students can select the correct answer by purely matching of words: there would be no need for faith and looking at a map are used in both the passages and the test items. This finding can be attributed to the teachers' lack of paraphrasing skill. As what Harris (1969) asserted, the correct answer of a test item should be paraphrased to avoid superficial testing of comprehension. Therefore, to better develop comprehension on the text, options should be paraphrased.

Moreover, 11 (19.29%) of the items are with options that do not have the same length. This implies that some options in the test items are obviously longer than the other ones. The following items show violation of that principle:

(RC7) As used in the beginning of the story, which best describes the act of being relegated?

- A. A star athlete is given a coveted award.
- B. A disobedient dog is put in the garage to sleep.
- C. After being tested, swimmers are put into groups according to their ability levels.
- D. The birthday boy is taken to his favorite restaurant.

In item RC7, option C is much longer than the other options. In the question of item RC7, students who neither read nor comprehend the reading text might simply choose option C because greater detail is provided. Students often have the notion that the option which has the greater detail is the correct answer, especially if the question is too difficult for them.

### Implication to Test Construction

Teachers assume several responsibilities aside from delivering instruction. With the multitude of tasks that teachers hold, assessment is sometimes compromised. With the analysis made in this study, it was found that the collected test papers have a great number of items that need improvement since they violate several principles in vocabulary, grammar and reading comprehension tests. The identified violations imply that there is a need of reviewing the rules and principles in constructing good and effective test items. Likewise, test developers should

acquire the knowledge in testing and apply all the necessary principles with regard to test construction so as to achieve what the test really measures. Students' exposure to well-constructed test items is a good practice for them to be engaged in higher order thinking, in preparation for taking entrance exams and board exams especially so that these tests are in multiple-choice type format. Moreover, the violations imply that constructing test is not easy, and that it must be taken seriously because a poorly constructed test could compromise the its validity and reliability. Therefore, one must exert effort and consider important principles when constructing test items. We teachers, as test developers, should be well prepared and knowledgeable before we construct test items and more importantly, we should write test items seriously because it may directly affect our students' academic performance and achievement. As what Allanson and Notar (2019) posited, educators have an obligation to ensure that the assessments they create accurately measure learning objectives.

### CONCLUSIONS

The present study determined the violated test construction principles reflected in selected summative test in English. The findings of the study reveal that most of the vocabulary items violated the principle on making all the options with the same word class. On the other hand, the majority of the grammar test items used mixed options. When it comes to the reading comprehension section, most of the items can be answered from general knowledge. In the light of the findings, it can be concluded that the test papers have violated varied principles in vocabulary, grammar and reading comprehension tests. It can also be concluded that several test items need improvement. Moreover, the violations reflected in the summative tests imply that these are a serious concern to test construction. Therefore, with the findings, conclusion drawn as well the study' limitations, several recommendations are offered. Teachers should review the principles of test construction and upgrade their test papers. They should continue attending to seminars which could help them improve their skills in constructing good and effective test items and developing valid and reliable testing materials. Likewise, their test papers should be checked and validated by experts in their respective schools before administering them to their students. Moreover, test construction should be included as a topic in their Learning Action Cell (LAC) session to address and in-service trainings. Since the study focused on multiple-choice only, it is recommended to investigate errors in other test types such as identification, fill in the blanks, matching type, true or false and essay type. This study determined the principles violated in terms of test item options, thus, it is also interesting to investigate test item errors in terms of stem. A similar study may likewise be conducted using a more test papers in English and other subject areas such as Mathematics and Science to further validate the findings of the study. Finally, the

contributions this study has made can be employed by teachers and test developers for the investigation of other violated test construction principles, using other models

## REFERENCES

- Allanson, P., & Notar, C. (2019). Writing multiple choice items that are reliable and valid. *American International Journal of Humanities and Social Science*, 5(3), 1-9. Retrieved from <https://ijehd.cgrd.org/images/vol5no3/1.pdf>
- Bachman, L., & Palmer, A. (1996). *Language testing and practice*. Hong Kong: OUP.
- Bailey, K. M. (1998). *Learning about language assessment: dilemmas, decisions, and directions*. Heinle & Heinle: US.
- Begum, T. (2012). A guideline on developing effective multiple-choice questions and construction of single best answer format. *Journal of Bangladesh College of Physicians and Surgeons*, 30(3), 159-166. doi: <https://doi.org/10.3329/jbcps.v30i3.12466>
- Brown, D. (2004). *Language assessment principle and classroom practice*. San Fransisco: Longman.
- Hansen, J. D., & Dexter, L. (1997). Quality multiple-choice test questions: Item-writing guidelines and an analysis of auditing test banks. *Journal of Education for Business*, 73, 94-97. doi: <https://doi.org/10.1080/08832329709601623>
- Harris, D.P. (1969). *Testing English as a Second Language*. McGraw-Hill, Inc, New York, U.S.A.
- Henning, G. (1982). Twenty common testing mistakes for EFL teachers to avoid. *English Teaching Forum*, 20 (3), 33-37. Retrieved from [https://americanenglish.state.gov/files/ae/resource\\_files/50\\_3\\_8\\_henning.pdf](https://americanenglish.state.gov/files/ae/resource_files/50_3_8_henning.pdf)
- Lin, Y. (2002). Principles for Writing Multiple-Choice Items in Vocabulary, Grammar, and Reading Tests: A Study on the Test Making Process of an Achievement Test at Fooyin Institute of Technology. Thesis. Fooyin Institute of Technology.
- Linn, R.L., & Gronlund, N.E. (1995). *Measurement and assessment in teaching*. Prentice-Hall, Inc, New Jersey, U.S.A.
- Madziyire, N. C. (2010). *Supervision and leadership*. Harare: Zimbabwe Open University.
- Mpofu, B. (2011). *Formative evaluation versus summative evaluation*. Harare: Longman.
- Rubin, D. L., Daly, J., McCroskey, J. C., & Mead, N. A. (1982). A review and critique of procedures for assessing speaking and listening skills among preschool through grade twelve students. *Communication Education*, 31(4), 285-303. doi: <https://doi.org/10.1080/03634528209384697>
- Sax, G. (2007). *Principles of educational and psychological measurement and evaluation*. Belmont: Wadsworth Publishing Company.
- Sharma, Mani. (2020). Perceptions on language testing and assessment: A Case study of B.Ed students' in Nepal. *Journal of Advances in Humanities and Social Sciences*, 6(1), 27-33. doi: 10.20474/jahss-6.1.4.