

Take the Test: An initial investigation on the common errors in multiple-choice questions in English quarterly assessments

Abstract – Classroom assessment is one of the most important aspects of the teaching-learning process. While there are several available references that offer excellent advice on preparation and construction of effective objective tests for students, there remains to be frequent reports of deficiencies or lapses in locally developed classroom examinations especially in multiple-choice tests. The aim of this descriptive study was to determine the common errors in multiple choice questions (MCQs) in teacher-made quarterly assessments in English in one public secondary school in Ilocos Norte, Philippines. Frequency counts and percentages were used to treat the data gathered. The interpretation and analysis of data utilized the basic principles of constructing MCQs used in the study of Kar, Lakshminarayanan, and Mahalakshmy (2015). The findings show that the multiple-choice questions in English quarterly assessments had errors related to content issues, writing the stem, writing the distractors, and formatting and style issues. Of these four categories of errors, the most committed errors were related to writing the stem while the least committed errors were related to content issues. It can be concluded that, generally, nearly half of test items in English quarterly assessments need to be reconstructed or corrected as these have negative implications to the students' test performance and, subsequently, to their academic standing. **Keywords** –multiple choice questions, test construction, distractors, assessment,

Comment [I1]: It covers all aspects of the paper but misses explicit reference to the research method (content analysis) and the tool selected (Checklist for Constructing Effective Multiple-choice Questions Items).

INTRODUCTION

The need to make a well-constructed summative test, in any subject area, more so in Language classes, is essentially imperative. Several research have proven that poorly constructed tests adversely affect the students' achievement in examinations. Meanwhile, the multiple-choice question (MCQ) is a common format used to assess students in English classes, particularly, in the assessment of separate components of language knowledge, like grammar, vocabulary, reading comprehension, among other skills. The two fundamental components of a standard multiple-choice test item include the problem or the stem and potential answers, also called alternatives (Burton, Sudweeks, Merrill, & Wood, 1991). Burton, Sudweeks, Merrill, and Wood (1991) also explained that the stem can be either a question or an unfinished statement and that there is one best or accurate alternative (answer) on the list of alternatives, as well as a number of less desirable or distracting options. Teachers, schools, and assessment organizations choose to employ multiple-choice questions (MCQs) to evaluate students' accomplishments because of its advantages like easy, quick, and economical scoring (Linn & Gronlund, 1995). Furthermore, Terrant and Ware (2008) noted that teachers can easily evaluate a large number of candidates using the MCQ format, and they can test a variety of subjects. They further underscored that if multiple-choice tests are appropriately designed, MCQs can evaluate higher levels of cognitive thinking and accurately distinguish between students with high and low achievement. In addition to this, Bontis (2009) highlighted that compared to other types of evaluations, multiple-choice tests are the best predictors of students' overall success because they are scored objectively and thus make the test more reliable than subjectively scored tests.

However, Terrant and Ware (2008) lamented that MCQs in many examinations developed in-house *are poorly constructed* [emphasis mine], and usually this is attributed to teachers' inadequate knowledge and preparation for creating well-designed assessments. Terrant and Ware

Comment [I2]: Too long paragraphs

(2008) also expressed disappointment over the fact that, despite multiple publications clearly outlining the rules for creating high-quality MCQs, violations of these rules are nevertheless frequent. Such observation is relevant in Philippine schools. In fact, the Department of Education, through its bureaus in charge of assessment considers this a serious matter. There have been Department orders and memoranda cascaded to the regional and division offices and schools along this concern. In one public secondary school in Ilocos Norte, Philippines, for example, it has been an observation by the head teachers and the principal, who are in-charge of checking quarterly assessments, that there are frequent commissions of errors in the tests prepared by teachers across subject areas. Most badly written multiple-choice test questions include at least one of the following three flaws, although not always all three: they attempt to measure an objective that is not well-suited for them, they include hints for the right answer, or they are ambiguously phrased. Notably, students' performance on MCQs may be impacted by errors in the item's writing, which may make answers more or less difficult. Again, as found in results of various studies along this area, flawed items can negatively affect student's academic achievement. Hence, studies that aimed at investigating teacher-made examinations for error-analysis has never been more imperative and compelling than today when quarterly assessments constitute a large part of the students' total grade.

Such investigations will also contribute to the existing data on the quality of teacher-made or locally developed tests used for assessment that will be utilized by Department of Education officials as bases in crafting policies on classroom evaluation and assessment.

It is in this light, that the researcher aimed to determine the common errors in the multiple-choice questions in English Quarterly Assessments in one of the public secondary schools in Ilocos Norte, Philippines. Specifically, this study aimed to seek answers to the question:

1. What are the errors in multiple-choice questions committed in the quarterly assessments in English that relate to:
 - 1.1 content issues
 - 1.2 writing the stem
 - 1.3 writing the distractors
 - 1.4 formatting and style issues

Literature Review

The teaching profession is a very flexible and fascinating one because it involves not only planning and executing classes but also evaluating students' performance and level of learning, hence, creating examinations and figuring out how to assess student performance are crucial components of education (Griese, 2005).

In the Philippine education system, classroom assessment is one of the most vital aspects of the teaching-learning process. It must be noted that it takes a lot of time, resources, and effort for teachers to create effective assessment tools because they are used to determine the learning outcomes. Given its significance in determining students' performance in the classroom, classroom assessment, most especially quarterly assessments should be done with due diligence. While there are several available references that offer excellent advice on preparation and construction of effective objective tests for students, there remains to be frequent reports of deficiencies or lapses in teacher-made classroom examinations. One of the commonly reported is related to poorly constructed multiple-choice test items. The use of multiple-choice questions in testing remains to be one of the most enduring and successful forms of educational assessment to date. MCQ-based tests that have been properly constructed are effective, objective,

Comment [13]: A long paragraph.

Comment [14]: No in-text citation though.

discrimination-capable, and can be integrated with other assessment methods to contribute to a holistic student assessment plan (Brady, 2005). In addition, the use of multiple-choice questions in educational assessments allows for the measuring of a variety of knowledge, skill, and competency levels. (Haladyna, 2004; Downing, 2006; Popham, 2008 as cited in Shin, Guo, & Gierl, 2019). Other advantages of multiple-choice items include but not limited to being efficient to administer; they are easy to score objectively; they can be used to sample a wide range of content; they require a relatively short time to administer (Haladyna, 2004; Haladyna and Rodriguez, 2013; Rodriguez, 2016 still in Shin, Guo, & Gierl, 2019). Hence, educational assessments experts warn subject teachers to construct multiple-choice tests based on existing principles in test construction because learning is likely to be distorted by poorly-written MCQs. Kar, Lakshminarayanan, and Mahalakshmy (2015) emphasized that the quality of MCQ depends on the quality of the items overall and the use of effective decoys. Furthermore, they made the point that poor MCQs interfere with the assessment process, making it crucial to create dependable and valid items that are devoid of these faults.

Given these reports and observations on faulty multiple-choice questions, Downing (2005) in his study, suggested that efforts to teach faculty the principles of effective objective-test item writing should be increased. More importantly, it must always be noted that MCQs that are of high quality, dependable, and capable of testing high order thinking abilities, and congruent with curricular objectives, a significant amount of time, effort, and commitment is needed (Kar, Lakshminarayanan, & Mahalakshmy, 2015).

Comment [15]: Only 5 references included in this section, mostly not updated.

Materials and Methods

To attain the objective of the study which is to determine the common errors in multiple-choice questions in English quarterly assessments, and to answer the research question, a total of 18 teacher-made test papers used for Quarterly Assessment in English from various grade levels (Grade 7-11) were randomly selected for analysis. From these 18 test papers, only the multiple-choice questions (test items) were gathered, and the other test types were disregarded. A total of 352 test items composed of items for vocabulary, grammar, reading comprehension, and oral communication skills were subjected to content analysis to identify the errors in test items. While there are several principles, guidelines, and standards in constructing multiple choice test items set by various educationists and researchers, this study opted to adopt the *Checklist for Constructing Effective Multiple-choice Question Items* used in the study of Kar, Lakshminarayanan, and Mahalakshmy (2015) to gather the data on multiple-choice test construction errors because it is more convenient to use, easier to understand, and it is more organized, encompassing, and holistic. Using the said checklist, during the analysis for errors in MCQs, there were four categories that were investigated, namely: content issues, writing the stem, writing the choices/distractors, and formatting and style issues. Under each category were specific guidelines or principles descriptive of each category. However, the set of principles were not used in its entirety and only those which were deemed necessary and relevant to the study were used. Under content issues, the principles include 1) base each item on important content area / learning outcome; 2) ensure each item is wholly independent of the others. Under the category writing the stem, the principles are: 1) ensure that the directions in the stem are very clear; 2) include the central idea and common elements in the stem; 3. keep it clear, concise and unambiguous; and, 4) keep vocabulary simple. The principles that are under writing the distractors include: 1) keep choices independent/ mutually exclusive; 2) make sure that only one of these choices is the right answer; 3) keep choices homogeneous in content (distractors are in

the same category as the correct answer); 4) make all distractors plausible, arrange choices in logical or numerical order, wherever appropriate; 5) and keep the length of choices about equal. Under the formatting and style issues, the following principles are to be observed: 1) all parts of the item should be in the same page; 2. use correct grammar, punctuation, capitalization, and spelling; and, 4) keep it brief and minimize the amount of reading in each item. Frequency counts and percentage distribution were used to treat the data which were later analysed and interpreted.

RESULTS AND DISCUSSION

This part presents the data gathered and the interpretation of the results of the analyses made. The discussions of the findings are arranged according to the four categories of errors in multiple-choice questions: errors in multiple choice questions related to content issues; errors related to writing the stem; errors related to writing the distractors; and errors related to formatting and style issues. The last table shows the distribution of errors in each category.

Errors Related to Content Issues

Table 1 shows the number of errors related to content issues. The table further indicates that of the 352 test items, there were only three (3) or .85% that violated that principle, *Ensure each item wholly independent of the others*. This means that there were three questions that were related or connected to the preceding or succeeding questions. In principle, multiple-choice test constructor needs to ensure that each item can stand alone by itself or is independent of the other items in the test. For example, in the First Quarterly Examination in English 9, to answer item number 15, test-takers must base it in item number 14: *What does the line in item number 14 mean?* Such violation of the said principle may lead to students' confusion and inconvenience in taking the examination.

Principle	Frequency (f)	Percentage
Base each item on important content area / learning outcome	0	0
Ensure each item is wholly independent of the others	3	.85%
Sub-total	3	.85%

Table 1. Content-related errors in multiple choice test items

It can also be gleaned from the table that there was no (0) error committed as regards the rule, *Base each item on important content area/learning outcome*. This means that there were no questions on subject matters that were not taught to the students in the classroom for the particular grading period.

Errors Related to Writing the Stem

Table 2 shows errors in multiple-choice questions related to writing the stem. It further reveals that there were 63 items found to be erroneous in terms of writing the stem in MCQs. Of this number, there were 35 (9.94%) items that showed violations of the principle, *keep the stem clear, concise, and unambiguous*. It has always been clearly underlined in several publications that test items must not be wordy, tricky, and unjustifiably lengthy, however, in the study, ironically, it was found that this was the most violated principle.

Table 2. Errors related to stem writing.

Principle	Frequency (f)	Percentage
Ensure the directions in the stem are very clear.	14	3.98%
Include the central idea and common elements in the stem	13	3.69%
Keep it clear, concise, and unambiguous	35	9.94%
Keep the vocabulary simple	1	.28%
Sub-total	63	17.89%

For example, is the item: *All the following is an example of small group communication, except one, which of these is not?* Aside from the grammatical lapse in terms of subject-verb agreement committed by the test constructor in the test item, it can be noticed from the item that it is syntactically ambiguous because of the word order used. Such item can confuse students which might lead them to get the wrong answer and will eventually result to having low test score. Another test item that demonstrates violation of the said principle is the item, *When talking to other people (small group), what do you need most?* Clearly, there is a gap in the item as it does not establish or clarify the context. Such vagueness in test items affects the student's comprehension of the question. Another item that violates the aforementioned principle is the item, *The Speech Act Theory was a concept proposed by _____*. This item displays another case of wordiness. The phrase "a concept" should have been omitted as it is deemed no longer necessary. Such unclear, unconcise, and ambiguous multiple-choice test items confuse students which eventually take much of their time answering the said items. In fact, Haldyana and Downing (1989, as cited in Brame, 2013) warned that the stem should not contain irrelevant material, which in effect, can decrease the reliability and the validity of the test scores.

Meanwhile, there were 14 (3.98%) of the 352 items that violated the principle, *Ensure the directions in the stem are very clear*. One item is written in this way: *Follow stories about speeches usually emphasize one of three things:*. Aside from the construction is ambiguous, it does not clearly tell what it exactly wants to elicit. Another item violative of this is rule is the item, *Saying a Lord's Prayer is an example of speech style*. The item, clearly, fails to specify what the question is. Such construction is detrimental to student's test performance. If directions in the stem are not clear, the test item itself becomes invalid, consequently, students will have a hard time answering the item.

On the other hand, 13 items or 3.69% of the 352 multiple choice questions showed errors in terms of the principle, *Include the central idea and common elements in the stem*. A test item that exemplified this error is the item:

The speech act theory is defined as:

- A. The theory that words are things
- B. The theory that each word has a speech or illocutionary act
- C. The theory that if you act in a certain way, then your speech will be perceived differently
- D. The theory that if you act in a certain way, then your speech will be perceived differently.

In the said item, aside from there is duplication of choices, C and D, the options has also common element, that is, the phrase, *The theory that*, which should have been included in the stem to lessen the time of reading the answers, hence, the students will have more time for the other test items. One more example of this violation is demonstrated in the test item below:

Complete the following sentence: Speech act theory helps to define the relationship of _____.

- A. one who is performing and one who is acting
- B. one who is speaking and one who is listening
- C. one who is being and one who is conceptualizing
- D. one who is listening and one who is creating a picture of the desired act

The common elements in the alternatives should have been included in the stem. Reading the choices in the item above is time-consuming. Another item that violated this rule is the test item, *A kenning is a _____*. The item lacks the central idea. It should have been rephrased so that the stem would be more meaningful and that it would not confuse the students. There was only 1 (.28%) item that violated the rule, *Keep the vocabulary simple*. In other words, almost all the test questions were understandable in terms of the level of vocabulary used.

Errors related to writing the choices/distractors

Table 3 shows that there are 39 MCQs that committed errors related to writing the choices or distractors. In terms of errors related to the writing of choices or distractors, the principles, *Keep choices homogeneous* and *Make all choices plausible* appeared to be the most violated, both of which, were committed in 12 items or 3.41% of the 352 test items. As regards the principle, *Keep choices homogeneous*, there were items that contained distractors that did not belong to the same category or class.

Table 3. Errors related to writing the choices/distractors.

Principle	Frequency (f)	Percentage
Keep choices independent/mutually exclusive	2	.57%
Make sure that only one of the choices is the right answer	0	0%
Keep choices homogeneous	12	3.41%
Make all distractor plausible	12	3.41%
Arrange choices in logical or numerical order, wherever appropriate	3	.85%
Keep the length of choices about equal	10	2.84%
Sub-Total	39	11.08%

For example, in the vocabulary test item below, the students were asked to give the meaning of the underlined words in the sentence.

All the students were interested in the lesson, so they actively took-part in the lively discussion.

- A. participated
- B. listened
- C. recorded
- D. take down

It can be gleaned from the item that the underlined idiomatic phrase is in the past tense, so technically, the answer must also be in the past tense. All options are in the past tense, except letter D which gives hint to students that it is automatically a wrong option. Another test item that violates this rule is the item asking the students to give the correct meaning of the underlined word in the sentence:

I am a great aficionado of rock music.

- A. skillfull
- B. clever
- C. unselfish
- D. follower

It can be observed in the item that the underlined word belongs to the word class, noun, so it expects a noun answer. Of the four choices, only D is a noun, the rest of the options are adjective. Given the options, the students are inclined to take D as the correct answer. There is another item asking the students to give the meaning of the underlined word in the sentence that violates the rule:

- I felt like an alien in my son's classroom because I was so much taller than the children.
- A. wealthy
 - B. unworthy
 - C. foreigner
 - D. poor

It can be noted in the test item that the underlined word is *alien* which functions as a noun in the sentence, hence, it expects that students will pick a noun as the answer. All the options belong to the word class, adjective, except option C (foreigner) which functions as a noun. Considering this premise, the item may be considered a give-away because of the hint indicated in the options. It has been highlighted in several references that alternatives should be homogenous in content. Alternatives that are heterogeneous in content can provide cues to student about the correct answer (Brame, 2013). Meanwhile, the items that violated the principle, *Make all distractors plausible* include the test items below:

- Damian L. Eugenio, the mother of Philippine folklore compiled collections of proverbs in our country.
- A. gathered and arranged
 - B. wrote and published
 - C. gather and arrange
 - D. write and publish

It can be noted in the test item that options C and D are no longer plausible since these are verbs stated in the present tense. The underlined verb is in the past tense, hence, it expects to have an answer in the present tense, too. The same principle is violated in the item:

- The Spaniards were able to interact with the early Filipinos because they learned their indigenous languages.
- A. foreign
 - B. acquire
 - C. native
 - D. intrinsic

In the item, option B is no longer plausible since it is obvious it is not of the same class as the other options. The test item below also consists of distractors which are not plausible:

- What three senses do the four lines appeal to?
- A. sight, smell, and taste
 - B. sight, touch, and hear
 - C. sight and taste
 - D. touch and hear

In the test item, the question clearly asks for the three senses, hence, options C and D are automatically no longer plausible since these options contain only two senses. It has been a rule that all alternatives should be plausible. In fact, Brame (2013) explained that the purpose of the erroneous alternatives is to act as distractions; students who did not meet the learning outcome should choose them, while those who did should ignore them. Implausible alternatives should not be used because they do not operate as functional distractors. The most reliable cause of distractions is typical student blunders.

On the one hand, there were 10 items or 2.84% of the 352 questions that violated the principle, *Keep the length of the choices about equal*. A test item that commits a violation to this rule is the MCQ item below:

When conducting a question-and-answer session, the following guidelines are appropriate except:

- A. answer as briefly as possible
- B. listen for the substance or big idea of the question
- C. paraphrase confusing questions before answering
- D. if the questioner is attacking you personally, use a subtle attack rather than give a defensive answer.

It can be observed from the item that option D is very much lengthier than the other options. At times, students think that the longest answer is the right answer. It is suggested that options must be of equal length to avoid giving the students a hint of the possible correct answer. There were 3 (.85%) items that violated the principle, *Arrange choices in logical or numerical order, wherever appropriate*. This violation is evident in the test item below:

The poet associated death to the sea. What is the reason for this?

- A. The vastness and depth of the sea is just like death is unknown and uncertain
- B. The sea is calm and so is death
- C. It is believed that in the olden times that the sea is the way to the underworld.
- D. Death is just like sailing away to a new dimension.

It can be observed from the test item that the options appear to be random in terms of length. Experts in test constructions suggest that the options must be arranged from the shortest response to the longest because aside from it facilitates better reading, and it is also more pleasing to the eyes of the test takers as the arrangement becomes more organized. There were 2 items that showed violation on the principle, *Keep the choices independent/mutually exclusive*. These were the items that included *All of the above and none of the above as options*. The inclusion of such options in MCQs is discouraged as these are immediately deemed by the students to be the right answer without them reading the whole question.

Errors related to formatting and style issues

Table 4 shows that under the category, formatting and style issues, the principle, *Use correct grammar, punctuation, capitalization, and spelling*, recorded the most number of violations with 33 items or 9.36% of the total number of MCQs.

Table 4. Errors related to formatting and style issues

Principle	Frequency (f)	Percentage
-----------	---------------	------------

All parts of the stem should be in the same page	3	.85%
Use correct grammar, punctuation, capitalization, and spelling	33	9.38%
Keep it brief and minimize the amount of reading in each item	2	.57%
Sub-total	38	10.8%

The test item, *These are sequential visual art, usually with text that are often told in a series of rectangular panels*, contains grammar lapse. It can be observed that in the test item, the word *art* should be in plural form, *arts*, as signaled by the phrase *These are ...*. Another item that has lapse in grammar is the item, *In the medieval era, wealthy knights won a woman's hand through bravery while troubadours, won them through _____*. The said test item violates the rule on noun-antecedent agreement. In the sentence, the noun, *hand* received *them*, and it must be *it*, since it is singular. Also, the test item, *Yuan ask Zhen to put the trash at the back of the school. Zhen said yes and hurriedly went to throw the trash. Zhen's action is considered as _____* indicates a lapse in grammar, specifically, it exemplifies errors on verb tenses and agreement. Yuan is a singular noun, hence the verb must be *asks*. Another test item violative of the said principle is the item, *This is the effect produced on the listener when they listen to a locutionary act or the so-called heard response*. It can be noted in that there is a problem in the item in terms of the agreement of the pronoun and its antecedent. In the item, the antecedent, *listener*, should have *he/she* as the pronoun. It must be recalled that when test items are flawed with grammatical errors, or faults in terms of punctuation, capitalization, and spelling, students might be hard up getting the correct answer as there are high chances of them being confused and misled.

Three (3) items or .85% of the total MCQs violated the rule, *All parts of the item should be in the same page*. The 3 items were usually written at the bottom part of the test papers, and because of absence of enough space, some options were put on the succeeding page. When this happens, students are burdened as they inconveniently browse the pages of the test papers to look for the correct answer which eventually brings confusion to them aside from it consumes much of their time. Only 2 items (.57%) violated the principle, *Keep it brief*. These were the items that were too long and verbose. It has always been discouraged for teachers to construct unjustifiably lengthy responses as these consume much students' time and add confusion as well.

Table 5. Distribution of total errors in multiple-choice questions in the four categories

Categories	Frequency (f)	Percentage
Errors related to content issues	3	0.85%
Errors related to writing the stem	63	17.89%
Errors related to writing the distractors	39	11.08%
Errors related to formatting and style issues	38	10.8%
Overall Total	143	40.62%

The table shows the total of errors in multiple-choice questions in each category: errors related to content issues, errors related to writing the stem, errors related to the distractors, and errors related to formatting and style errors; and the overall total number of errors committed in the multiple choice questions in English quarterly assessments. It can be gleaned from the table that of the four (4) categories of errors in multiple-choice test construction, the most committed errors are those that relate to writing the stem of the question which has a total of 63 or 17.89% of the gathered multiple-choice items. The least committed errors are related to content issues which have a total of 3 or .85% of the multiple-choice questions. The table further indicates that

there are 143 (40.62%) multiple-choice questions that violated the principles in constructing multiple-choice questions. The results pose alarm since nearly half of the multiple-choice test items gathered for this study are erroneous. Such commission of errors in multiple-choice tests in English subjects can adversely affect their test performance, much more that these items were used in quarterly assessments in English which constitute 20-30% (depending on their grade level) of student's overall grade based on existing DepEd orders. Hence, it is important that teachers construct multiple-choice questions, or any test item in general, with much caution, due diligence and total adherence to existing principles in test construction. In fact, Allanson, and Notar (2019) pointed out that educators have obligation to ensure accuracy in assessments to be able to measure properly whether or not the students attained the learning objectives.

CONCLUSIONS

The present study determined the errors in multiple-choice questions in English quarterly assessments based on the principles of constructing multiple choice questions forwarded by Kar, Lakshminarayanan, and Mahalakshmy (2015). The findings show that the multiple-choice questions had errors related to content issues, writing the stem, writing the distractors, and formatting and style errors. Of these four categories of errors, the most committed errors were related to writing the stem while the least committed errors were related to content issues. It can also be concluded that, generally, nearly half of test items in English quarterly assessments need to be reconstructed or corrected as these have negative implications to students' performance in quarterly assessments, and subsequently, in their academic standing. Based on the findings and the conclusions, the following recommendations are forwarded: first, teachers who are most of the time are the test developers or constructors of school-based quarterly assessments, must have a regular review of the principles in constructing multiple-choice questions and make sure to adhere to such principles when making tests. Second, school heads in-charge of checking test papers must carefully validate each item for possible violations or errors before the tests are administered to students. Third, school-based seminars and trainings must integrate topics on effective construction of MCQs. The present study is limited only to investigation on the errors in multiple-choice questions in English, hence, it is recommended that other researchers investigate errors in construction of other test types, and expand the investigation to other subject areas like Science, Mathematics, and Filipino, among others, using other approaches, methodologies and principles of test construction.

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>

Brady, A. (2005). Assessment of learning with multiple-choice questions. *Nurse education in practice*, 5 4, 238-42.

Brame, C., (2013) Writing good multiple choice test questions. Retrieved from <https://cft.vanderbilt.edu/guides-sub-pages/writing-good-multiple-choice-test-questions/>.

Comment [16]: Some other updated references could have been included. Generally, references are not enough updated, only 4 within the last 10 years, and 2 within last 5 years. One reference is not in the text: Brown, G., & Abdulnabi, H. (2017) and viceversa: Bontis, 2009 and Downing, 2005 are two in-text citations that have not been included in the references.

Brown, G., & Abdulnabi, H. (2017). Evaluating the quality of higher education instructor-constructed multiple-choice tests: Impact on student grades. *Frontiers in Education*, 2. <https://doi.org/10.3389/feduc.2017.00024>.

Comment [17]: Not in the text.

Burton, S.J., Sudweeks, R.R., Merrill, P.F., & Wood, B. (1991). How to Prepare Better Multiple-Choice Test Items: Guidelines for University Faculty .Retrieved December 1, 2018, from <https://testing.byu.edu/handbooks/betteritems.pdf>

Griese, E. (2005). Multiple choice - a useful testing method for teaching English as a foreign language. Munich, GRIN Verlag. Retrieved from <https://www.grin.com/document/64363>

Kar, S.S., Lakshminarayanan, S. & Mahalakshmy, T. (2015). Basic principles of constructing multiple choice questions. *Indian Journal of Community and Family Medicine*, 1(2), 65- 69. <http://www.researchgate.net/publication/304349654>.

Linn, R.L. & Gronlund, N.E. (1995). *Measurement and assessment in teaching*. Prentice-Hall, Inc, New Jersey, U.S.A

Shin, J., Guo, Q., & Gierl, M. J. (2019). Multiple-Choice Item Distractor Development Using Topic Modeling Approaches. *Frontiers in psychology*, 10, 825. <https://doi.org/10.3389/fpsyg.2019.00825>

Terrant, M. & Ware, J. (2008). Impact of item-writing flaws in multiple-choice questions on student achievement in high-stakes nursing assessments. Retrieved on December 1, 2018, from <https://nursing.ceconnection.com/ovidfiles/01709760-201303000-00002.pdf>