The Art of Questioning Used in CED Departmental Examination: An Analysis

Lynn M. Besa[*], Richard E. Parcon
[1]Rizal Technological University
[2, 3]Tañong High School, DepEd Marikina, Rizal Technological University
[*]Email: lmbesa02@rtu.edu.ph

Abstract

Assessment plays a vital role in the development and improvement of quality education. In addition, assessment refers to anything that diagnoses the student's learning and monitors their level of understanding. It is also considered to be the key to evaluating the effectiveness of teaching strategies (Hans, 2015). Traditionally, the most common way to measure achievement and proficiency in language learning has been the paper and pen test. However, various forms of assessment and the design of classroom tests have also undergone some changes in line with the latest development in foreign language teaching specifically the communicative language teaching method.

At Rizal Technological University, the College of Education devised a departmental examination program. This program is initiated to determine the capacity of the learners to answer test-item questions patterned after the licensure examination for teachers which aims to eventually prepare them in taking the said examination. Utilizing the quantitative descriptive method, the researchers tried to analyze the levels of questions used in the Departmental Examination. Specifically, this paper sought to identify the types of questions given in the CED departmental examination, determine the level of questions given in the CED departmental examination in terms of the Revised Blooms Taxonomy and Webb's Depth Knowledge (DOK) model and evaluate the overall distribution of cognitive rigor present in the CED departmental examination.

Keywords: Assessment, departmental exam, effectiveness, multiple choice test, questions

Introduction

Assessment plays a vital role in the development and improvement of quality education. In addition, Assessment refers to anything that diagnoses the student's learning and monitors their level of understanding by checking the learners' knowledge, understanding, strengths, and problems with language skills like listening, speaking, reading, writing vocabulary, and literary skills. In addition, it is also considered the key to evaluating the effectiveness of teaching strategies employed by language teachers (Hans, 2015). Traditionally, the most common way to measure achievement and proficiency in language learning has been the paper and pen test. Even though alternative forms of assessment are growing and becoming more popular with many language teachers, some are still using traditional assessment (Frank, 2012). The report of Teacher Education through School-Based Support in India (TESS), (2015), said that in assessing language learning in a regular classroom there are several ways that a language teacher can use to assess his/her students during normal class hours like combining vocabulary comprehension test and grammatical structure of a sentence, movie review, listening comprehension test, and others. It also pointed out that teachers do not need to conduct extra exercises or activities to fully assess the language proficiency of the students. Instead, utilizing the same routine that varies in instructions and level of difficulty will be enough to assess the learner's language proficiency.

Concerning the study of language assessment literacy, it is indeed a very important matter to attend to for all language teachers. Though in the depth of the literature review in the study, it was emphasized that language assessment literacy is a very critical field for researchers and other scho-

lars who would pursue to study this concept to add additional baseline data in language education specifically in language testing. Familiarizing and understanding the basics of validity and reliability of the test and the ability to perform test-related activities become more significant in a test-oriented situation. Also, Uncharted Area for the English Language Teachers in Bangladesh provided insights into how inadequate academic and professional testing background of teachers hindered their performance in conducting assessment-related tasks and gave limitations in utilizing assessment to improve teaching and learning. In the end, constant professional development opportunities are needed for the teachers to be knowledgeable and skilled in language testing activities (Sultana, 2019).

In addition to the various forms of assessment, the design of the classroom tests has also undergone some changes in line with the latest development in foreign language teaching specifically the communicative language teaching method. However, regardless of various perspectives and shifts in the field of language testing. Multiple choice type of test is still relevant and used by several language teachers to measure the language proficiency of learners. It is also emphasized in the study the analysis of multiple-choice items aimed at testing grammar, vocabulary, and reading comprehension. Findings revealed that most of the items are at a moderate level in terms of item facility. Besides, the result shows 28% of the items have a low item discrimination value. Lastly, the frequency results were analyzed in terms of distractor efficiency and it has been found that some distractors in the exam are significantly ineffective and should be revised (Toksoz & Ertunc, 2017).

In the article analyzing multiple choice type of tests, it was suggested that multiple choice type test items provide overloaded educators with the opportunity of an easy and consistent process of scoring and grading. It also maintained that multiple-choice items are easy to prepare because there is a computer program designed to prepare multiple-choice items using vocabulary (Brown, 2004). While tests play an important role in giving feedback to teachers on their educational actions, therefore the quality of a test is a critical issue (Quiragrain and Arhin, 2017). Moreover, a typical multiple-choice type of test consists of a question or an incomplete statement which is referred to as the stem, and a set of three to four options with one correct answer. The learners are asked to select one response for the question. The rest of the options are called distractors. Several articles stated that constructing a good distractor is very difficult. Therefore, constructing a multiple-choice type of test is not just a simple process, rather it is considered a complex process because of searching for plausible distractors. The easiest way of checking makes multiple-choice tests appealing to all the teachers who teach a large number of students in a language class. Another advantage is that a well-constructed multiple-choice test can yield test scores at least as reliable as those produced by a constructed-response rest, while also allowing for a large portion of the topics covered in a course to be assessed in a short period (Oluseyi & Olufemi, 2011).

In the study on evaluating the cognitive levels of master class textbook questions, it was emphasized that 52 percent of the questions measured the comprehension level of the learners while other questions in the textbook were categorized under higher-order thinking skills (Assaly & Smadi, 2015). While developing higher critical thinking skills is a must in every classroom, therefore EFL teachers should be experts in using various strategies in formulating and asking questions to the learners (Feng, 2013). Relative to analyzing the reading questions of textbooks in Jordan, it was revealed that low-level questions were frequently used in three textbooks, and at the university level, the textbooks do not measure a higher level of thinking but the reading content in secondary school textbooks focused on higher level questions. Therefore, communication between the secondary and university level should be done to address the gap between them (Freahat & Smadi, 2014). While, a content analysis was conducted on the WH-Questions in the EFL textbooks of Horizon, and found that most of the questions measured the high-level thinking of the learners and became an aid in de-

veloping cognitive skills (Igbaria, 2013). In addition, the cognitive level of questions in social studies textbooks and the views of teachers based on blooms taxonomy revealed that questions in the textbooks were not distributed well under the level of comprehension of bloom's taxonomy because low-level questions were dominant than higher-level questions (Tarman & Kuran, 2015).

When it comes to evaluating the learning objectives of textbooks in Iranian High Schools and Pre-University textbooks, the results showed that all the lower-order cognitive skills were more observed and measured than higher-order cognitive skills. In addition, the pre-university textbooks objectives utilized higher-order level skills which brought an implication to teaching and material development (Riazi & Mosalanejad, 2010). While the cognitive level of questions in social studies textbooks revealed that preparatory questions were lower in general compared to the cognitive levels of the students. Moreover, the assessment of teachers who were using the textbooks stated that the evaluation activity in textbooks utilized higher-order thinking questions (Tarman & Kuran, 2015). The WH questions in the English test were also evaluated, finding revealed that questions were aligned in the three-lower level of bloom's taxonomy and the results indicated that textbooks were not able to engage learners in the questions requiring higher levels of cognitive learning (Takasana, 2021). On the other hand, question analysis in textbooks are still many and the author utilized some questions that only require answer without justification and only needs a single procedure (Raditya et al., 2020).

At Rizal Technological University, the College of Education devised a departmental examination program. This program was initiated by the college to determine the capacity of the learners to answer test-item questions patterned after the licensure examination for teachers which aims to eventually prepare the teacher education students in taking the said examination. Though the questions were constructed by the assigned faculty to teach the Professional Education subjects, questions were not fully validated if they measure the intended outcome of the course and the level of thinking skills of the learners. This prompted the researchers to analyze levels of questions used in the Departmental Examination specifically it sought to answer the following questions:

1. What are the types of questions given in the CED departmental examination?
2. What is the level of questions given in the CED departmental examination in terms of the Revised Blooms Taxonomy and Webb's Depth Knowledge (DOK) model?
3. What is the overall distribution of cognitive rigor present in the CED departmental examination?

The result of this study will help the instructors better understand which level of test questions are highlighted in the departmental examination in developing learners' thinking skills. Also, it will help the teacher education students familiarize themselves with the type of questions used in the licensure examination for teachers. Lastly, the result of this study will help the College of Education to conceptualize a specific program that will enhance the skills of the teachers in writing test items for departmental examinations.

Since the focus of this study is to analyze the levels of questions used in the examination, this research is anchored with the revised Bloom's taxonomy and Webb's Depth Knowledge (DOK) model. From the concept of Benjamin Bloom about the taxonomy of learning, Anderson and Krathwohl (2001) came up with a new version of Bloom's taxonomy at the beginning of the 21st century. The following terms were included in the taxonomies of the cognitive domain proposed by Anderson and Krathwohl (2001), *Remembering;* is the process of recognizing and recalling knowledge from memory. It is also used to produce or retrieve definitions, facts, or lists or to recite previously learned information. *Understanding*, wherein the construction of meaning from different types of functions is written on various platforms. *Applying* carries out or uses procedure through executing

or implementing. It also relates to situations where learned material is used through products. *Analyzing*; breaking materials or concepts into parts, determining how the parts relate to one another, how they interrelate, or how they relate to an overall structure or purpose. ***Evaluating***; making a judgment based on criteria and standards through checking and critiquing. Lastly, *Creating* wherein it combines elements to form a coherent or functional whole, recognizing elements into a new pattern or structure through generating, planning, or producing.

Webb's Depth of Knowledge (DOK) model ensures that teachers teach to a level that will promote student achievement. In terms of assessment, this model is considered as a mechanism to ensure that the intended course outcome and the level of students' demonstration matches the assessment employed across the curriculum. DOK Level 1 deals with recall and reproduction, such as basic recall of concepts, definitions, facts, and processes. In assessment, learners may answer Level 1 questions that involve following simple procedures or formulas. DOK level 2 focuses on the primary application of skills and concepts**,** such as engagement of some mental processing beyond recalling or reproducing a response. Test item questions require the learners make some decisions as to approach the question or problem. DOK level 3 emphasizes strategic thinking, which requires a deep understanding of the concept as reflected in planning, using evidence, and more demanding cognitive reasoning. In assessment, a test item question may have more than one possible answer, and learners need to justify it. Lastly, the DOK level 4, which is the highest level of depth of knowledge, requires a high cognitive demand, and it is very complex. Assessment is not used in multiple-choice tests because it requires a period to attain this level (Hess, 2004).

Most of the literature agreed that the revised Bloom's taxonomy is the most helpful model for categorizing questions according to their cognitive complexity. In comparison, the DOK model is also valuable for determining the depth of students' comprehension and complexity of the content in the required task and the type of questions used in the assessment. The combination of these two models helped the researchers to analyze the test questions in the departmental examination holistically.

Materials and Methods

Methods

This study utilized a qualitative research design to explore and understand the phenomenon (Creswell, 2019). Particularly, content analysis was used to analyze and categorize the recorded data for classification, summarizing, and tabulation

Sources of data

The data utilized were the test items from the departmental examination for Professional Education Subjects offered in the last 2nd semester of the school year 2021-2022.

The table below shows the distribution of test items according to the Professional Education courses and the number of test items written.

Table 1. Distribution of Test Items

Professional Education courses	Number of test items	Percentage
Prof Ed 11- Assessment of Learning 2	100	16.13
Prof Ed 9- Building and Enhancing New Literacies Across the Curriculum	60	9.68
Prof Ed 7- Technology for Teaching and Learning 1	70	11.29
Prof Ed 6- The Teacher and the Curriculum	120	19.35

Professional Education courses	Number of test items	Percentage
Prof Ed 5- The Teacher and the Community, School Culture and Organizational Leadership with a focus on the Philippine Technical Vocational Education and Training (TVET) System	100	16.13
Prof Ed 4- Facilitating Learner-Centered Teaching	70	11.29
Prof Ed 3- Foundation of Special and Inclusive Education	100	16.13
TOTAL	**620**	**100**

As reflected in this table, 6 professional education courses were offered during the second semester of the school year 2021-2022 namely Building and Enhancing New Literacies Across the Curriculum (Prof Ed 9) with 60 or 9.68 percent of the item's test questions. Technology for Teaching and Learning 1 (Prof Ed7) with 70 or 11. 29 percent of the test item questions. The Teacher and the Community, School Culture, and Organizational Leadership with a focus on the Philippine Technical Vocational Education and Training (TVET) System (Prof Ed 4) with 100 or 16.13 percent of the test item questions. Facilitating Learner-Centered Teaching (Prof Ed 4) with 70 or 11.29 percent of the test item. Foundation of Special and Inclusive Education (Prof Ed 3) with 100 or 16.13 percent of the test item questions in the departmental examination. A total of 620-item test questions were administered during the departmental examination.

Data Collection and Analysis

The data collection was divided into two stages. This first stage was the collection of test materials and the second stage was the categorization of the questions according to Bloom's Taxonomy and Webb's Depth of Knowledge or DOK (model) and the determination of the frequency of test items. Once the data were categorized, they were presented in tabular form and analyzed and interpreted based on the theoretical model presented.

Results and Discussions
Types of questions given in the CED departmental examination
Multiple Choice Type of Test

Table 2.Type of Test

Professional Education Courses	Range of Questions	Type of Test
3	1-100	Multiple choice
4	1-70	Multiple choice
5	1-100	Multiple choice
6	1-120	Multiple choice
7	1-90	Multiple choice
9	1-60	Multiple choice
11	1-100	Multiple choice

Data shows that all the tests, or 100 % of the tests facilitated among the CED students during the Departmental exam, were accounted to be multiple choice type of tests. As seen in the data, the test items are examples of a single select multiple-choice question. The Robert Gillespie Academic Skills Centre identified common types of multiple-choice questions. It could be extracted from the

test items that the types of multiple-choice questions are classified as a look-alike set of options. This means that the given options in the multiple-choice type of tests look very similar. Moreover, it could also be gleaned that the items are examples of flipping the text of the definition with the term, which means that some of the test items are generally identifying the key term based on a definition. It could also be gleaned that the multiple-choice type of question contains multiple options, which means that there is more than one correct option. Notice that a number of test items reveal some degrees of change from a study. This illustrates that the topics are taken out of their usual context, and the test takers are expected to predict what will happen.

Level of questions given in the CED departmental examination in terms of:
a.   Revised Blooms Taxonomy, and;
b.   Webb's Depth Knowledge (DOK) model

Table 3. Level of questions given in the CED departmental examination in terms of Revised Blooms Taxonomy

PROF ED COURSES	REMEMBERING	%	UNDERSTANDING	%	APPLYING	%	ANALYZING	%	EVALUATING	%	CREATING	%	TOTAL	%
Prof Ed 3- Foundation of Special and Inclusive Education	33	33	28	28	14	14	8	8	12	12	5	5	100	100
Prof Ed 4- Facilitating Learner-Centered Teaching	14	20	32	45.7	9	13.3	7	10	8	11	0	0	70	99.6
Prof Ed 5- The Teacher and the Community, School Culture and Organizational Leadership with a focus on the Philippine Technical Vocational Education and Training (TVET) System	54	54	31	31	8	8	5	5	2	2	0	0	100	100
Prof Ed 6- The Teacher and the Curriculum	87	72.5	23	19.2	8	6.67	1	0.83	1	0.8	0	0	120	100
Prof Ed 7- Technology for Teaching and Learning 1	15	16.7	19	21.1	21	23.3	18	20	7	7.8	10	11.11	90	100
Prof Ed 9- Building and Enhancing New Literacies Across the Curriculum	23	38.3	22	36.7	13	21.7	2	3.33	0	0	0	0	60	100
Prof Ed 11- Assessment of Learning 2	17	17	25	25	16	16	24	24	17	17	1	1	100	100
Total	34.71	35.928	25.71	29.529	12.71	14.653	9.29	10.166	6.71	7.229	2.29	2.444		

Data shows that 35. 928 % of the test items are categorized under Remembering. 29.529% of the items are understanding. 14.653% of the items are applying. 10.166% are evaluating, and only 2.444% are under ng.

Results show that Remembering has the highest percentage in terms of the no. of items. On the other hand, creating contains the least no. of items. The result indicates that based on Blooms taxonomy of learning, test items are more on remembering. Thus, it could be gleaned that Remembering is being highlighted in the test questionnaires.

Level of questions given in the CED departmental examination in terms of Webb's Depth Knowledge (DOK) model

Table 4. Webb's Depth Knowledge (DOK) model

	LEVEL 1 (RECALL AND REPRODUCTION)	%	LEVEL 2 (SKILLS AND CONCEPTS)	%	LEVEL 3 (STRATEGIC THINKING/REASONING)	%	LEVEL 4 (EXTENDED THINKING)	%	TOTAL	%
Prof Ed 3- Foundation of Special and Inclusive Education	48	*48*	25	*25*	27	*27*	0	*0*	100	*100*
Prof Ed 4- Facilitating Learner-Centered Teaching	22	*31.4*	30	*42.9*	18	*25.7*	0	*0*	70	*100*
Prof Ed 5- The Teacher and the Community, School Culture and Organizational Leadership with a focus on the Philippine Technical Vocational Education and Training (TVET) System	79	*79*	13	*13*	8	*8*	0	*0*	100	*100*
Prof Ed 6- The Teacher and the Curriculum	106	*88.3*	13	*10.8*	1	*0.83*	0	*0*	120	*100*
Prof Ed 7- Technology for Teaching and Learning 1	30	*33.3*	30	*33.3*	30	*33.3*	0	*0*	90	*100*
Prof Ed 9- Building and Enhancing New Literacies Across the Curriculum	39	*65*	20	*33.3*	1	*1.67*	0	*0*	60	*100*
Prof Ed 11- Assessment of Learning 2	24	*24*	42	*42*	34	*34*	0	*0*	100	*100*
Total	49.714	*52.714*	24.714	*28.614*	17	*18.643*	0	*0*		

It could be gleaned from the table that based on Webb's Depth Knowledge (DOK) model, 52.714% of the test items are categorized as level 1 (Recall and Reproduction). 24.714% of the test items are under level 2 (skills and concepts). 18.643% of the items are under level 3 (Strategic think-

ing/ reasoning). However, data shows that 0% or no items are accounted for under level 4 (extended thinking).

Results reveal that more than half of the items are categorized under level 1 (recall and reproduction) questions. Data shows that using Webb's Depth Knowledge, test items are dominantly only under level 1 (recall and reproduction).

The overall distribution of cognitive rigor present in the CED departmental examination

Table 5. Revised Blooms of Taxonomy

BLOOMS REVISED TAXONOMY	NUMBER OF ITEMS	PERCENTAGE
REMEMBERING	243	37.97
UNDERSTANDING	180	28.13
APPLYING	89	13.91
ANALYZING	65	10.16
EVALUATING	47	7.34
CREATING	16	2.50
TOTAL	640	100

In the overall distribution of the Revised Blooms Taxonomy of learning, it could be gleaned that 242 total items, equivalent to 37.97%, are under the category Remembering. One hundred eighty items, or 28.13%, are under understanding. Eighty-nine of the items, or 13.91%, are under-applying. Sixty-five items, or 10.16%, are categorized as analyzing, and 16 items, or 2.50% of the total 640 items, are categorized under creating.

Findings show that based on the total no. of items, Remembering, which is the lowest level in the Revised Blooms Taxonomy, has most of the items. On the other hand, creating, which is the highest level in the Blooms Taxonomy, has the least no. of items.

Table 6. Webb's Depth Knowledge (DOK) model

DEPTH LEVEL OF KNOWLEDGE (DOK model)	number of items	Percentage
LEVEL 1 (RECALL AND REPRODUCTION)	348	54.38
LEVEL 2 (SKILLS AND CONCEPTS)	173	27.03
LEVEL 3 (STRATEGIC THINKING/REASONING)	119	18.59
LEVEL 4 (EXTENDED THINKING)	0	0
Total	**640**	**100.00**

It could be assessed that based on Webb's Depth Knowledge Model, 348 items out of the total 640 items, which are equivalent to 54.38%, are questions under Level 1 (Recall and reproduction). One hundred seventy-three items, or 27.03%, are under level 2 (Skills and Concepts). One hundred nineteen items, or 18.59%, are under level 3 (Strategic Thinking/Reasoning), while no items under level 4 are accounted for.

Testing is one of the many proven ways of measuring the level of achievement or performance of learners. Multiple-choice type of test is one of the most commonly used types of test.

In most of the major exams, including the high-stake examination, testing centers typically use multiple-choice. The common notion is that they are easy to prepare and score. Contrary to this

notion, Burton et al. (1991) presented several reasons why some people are reluctant to use the multiple-choice type of test. Burton mentioned that some believe that these items are only good for measuring simple recall of facts. In fact, Multiple Choice Items are frequently used to measure lower–level objectives such as those based on knowledge of terms, facts, methods, and principles. However, the real value is their applicability in measuring higher–level objectives, such as those based on comprehension, application, and analysis.

Findings reveal that the College of Education in RTU used multiple-choice type of tests in the facility of its departmental examinations. It could be gleaned that this type of test was used as a form of summative assessment, which aimed to measure the academic achievement of college students, make decisions about their performance, and predict their possible performance during the actual licensure exam.

However, it could also be gleaned that most of the items set in the tests are categorized to be forms of remembering questions, the lowest level of cognitive learning under Blooms taxonomy of learning. Similarly, when the tests were subjected to Webb's Depth Knowledge (DOK) model, the same result was accounted for. The majority of the test items are categorized under level 1 (recall and reproduction). As such, the data illustrates that the multiple-choice type of tests which are single select multiple choice questions, classified as look-alike sets of options, flipping the text of the definition with the term, multiple options, and revealing some degrees of change from a study is designed to enhance the lower level thinking skills of the aspiring pre-service teachers.

Despite the many advantages of the multiple-choice type of test, it is often criticized. The findings of the study affirm the claim that Di Battista and Kurzawa in 2011 pointed out that multiple-choice items highlight what students can remember and do not assess the extent to which they can understand, apply and analyze course-related information (Walsh & Seldomridge, 2006). Although it is clear that thoughtfully written MC items can serve to assess higher-level cognitive processes, which require more skill than writing memory-based items (Buckles & Siegfried, 2006; Palmer & Devitt, 2007), the test items fell short of achieving higher-order thinking skills.

Conclusion

Multiple-choice type of tests is difficult to construct. Hence, creating well-written multiple-choice Items generally requires a more complex and time-consuming effort. The ability to write multiple–choice items as the type of test for a departmental exam is an important skill for the teacher to develop, particularly with the aim of preparing the learners for a higher and more high-stake exam such as the licensure examination. Equal distribution of the number of items under each level of cognitive rigor should be planned. Therefore, it is recommended that prior to the construction of the test items, curriculum mapping should be performed to project the coverage of the exam. Moreover, since the departmental exam aims to prepare the test takers to perform significantly in the licensure exam, the test makers should have rigid training on the competencies and scope of item constructions.

Acknowledgment

References

Assaly, I. R., & Smadi, O. M. (2015). Using Bloom's Taxonomy to Evaluate the Cognitive Levels of Master Class Textbook's Questions. *English Language Teaching*, *8*(5), 100-110. http://dx.doi.org/10.5539/elt.v8n5p100

Burton, J. S., [et al] (1991), How to Prepare Better Multiple-Choice Test Items: Guidelines for

University Faculty. Department of Instructional Science. Brigham Young University Testing Services. Web site, 16 August of 2009: http://testing.byu.edu/info/handbooks/betterItems.pdf

Brown, H D. (2004). Language Assessment: Principles and Classroom Practices. White Plains NY: Pearson Education.

DiBattista, David and Kurzawa, Laura (2011) "Examination of the Quality of Multiple-choice Items on Classroom Tests," The Canadian Journal for the Scholarship of Teaching and Learning: Vol. 2: Iss. 2, Article 4. DOI: http://dx.doi.org/10.5206/cjsotl-rcacea.2011.2.4 Available at: http://ir.lib.uwo.ca/cjsotl_rcacea/vol2/iss2/4

Feng, Z. (2013). Using Teacher Questions to Enhance EFL Students' Critical Thinking Ability. *Journal of Curriculum and Teaching*, *2*(2), 147-153. http://dx.doi.org/10.5430/jct.v2n2p147

Frank, J. (2012). The Roles of Assessment in Language Teaching. English Teaching Forum Vol.3 page 2.

Freahat, N. M., & Smadi, O. M. (2014). Lower-order and Higher-order Reading Questions in Secondary and University Level EFL Textbooks in Jordan. *Theory & Practice in Language Studies*, *4*(9), 1804-1813. http://dx.doi.org/10.4304/tpls.4.9.1804-1813.

Hans, Y. (2015). The Role of Assessment in Language Teaching, Learning, and Materials Development. https://www.dlsu.edu.ph/wp-content/uploads/pdf/conferences/research-congress-proceedings/2018/lli-01.pdf

Hess, K.K. (2004). Applying Webb's depth-of-knowledge (DOK) levels in reading. [online] http://www.nciea.org/publications/DOKreading_KH08.pdf

Igbaria, A. K. (2013). A Content Analysis of the WH-Questions in the EFL Textbook of" Horizons". *International Education Studies, 6*(7), 200-224. http://dx.doi.org/10.5539/ies.v6n7p200

Lorin W. Anderson And David R. Krathwohl. (2001). *A Taxonomy for Learning, Teaching, and Assessing*. https://www.uky.edu/~rsand1/china2018/texts/Anderson-Krathwohl%20-%20A%20taxonomy%20for%20learning%20teaching%20and%20assessing.pdf

Oluseyi A E., and Olufemi, A T. (2011). The Analysis of Multiple- Choice Items of the Test Scores in Chemistry in a Nigerian University. From Toksoz and Ertung (2017). Item Analysis of a Multiple-Choice Exam. *Advances in Language and Literary Studies Journal. Australian International Academic Center Publication*. http//dx.doi.org/10.7575/aiac.alls.v.8n.6p.140

Quaigrain, K and Arhin, A. (2017). Using reliability and item analysis to evaluate a teacher-developed test in educational measurement and evaluation. Cogent Education (2017), 4: 1301013.
https://www.tandfonline.com/doi/pdf/10.1080/2331186X.2017.1301013?needAccess=true

Raditya, A., Iskandar, R. S. F., & Suwarno, S. (2020). Questions Analysis in Mathematics Textbook from Competency-Based Curriculum up to Curriculum 2013. *Desimal: Jurnal Matematika*, *3*(2), 89-98. https://doi.org/10.24042/djm.v3i2.5973

Riazi, A. M., & Mosalanejad, N. (2010). Evaluation of Learning Objectives in Iranian High-School and Pre-University English Textbooks Using Bloom's Taxonomy. *TESL-EJ*, *13*(4). https://tesl-ej.org/pdf/ej52/a5.pdf

Sultana, N. (2019). *Language assessment literacy*: an uncharted area for the English language teachers in Bangladesh. https://link.springer.com/content/pdf/10.1186%2Fs40468-019-0077-8.pdf

Takasana, I. K. G. (2021). An evaluation of Wh-question in English in mind textbook a bright English textbook seen in terms of bloom's new taxonomy. *Journal of English Language and Literature Teaching*, *5*(2). https://doi.org/10.36412/jellt.v5i2.2454

Tarman, B., & Kuran, B. (2015). Examination of the cognitive level of questions in social studies textbooks and the views of teachers based on Bloom taxonomy. *Educational Sciences: Theory & Practice*, *15*(1). https://doi.org/10.12738/estp.2015.1.2625

Teacher Education School-Based Support in India (TESS). (2017). Supporting language learning through formative assessment. Teacher Education through School-based Support in India. https://www.open.edu/openlearncreate/pluginfile.php/134914/mod_resource/content/4/SE13_AIE_Final.pdf

The Robert Gillespie Academic Skills Centre. 10 Common Types of Multiple- Choice Questions and How to Solve Them. https://www.utm.utoronto.ca/asc/sites/files/asc/public/shared/pdf/tip_sheets_study/MultipleChoice_TypesQ_v10.pdf