Original Article

How Effectively are University Students Tested? A Case Study

Dr. Jane Kembo[1]*

[1] Rongo University, P. O. Box 103 – 40404, Rongo, Kenya.
[*] Correspondence email: jkembo@rongovarsity.ac.ke
 ORCID: https://orcid.org/0000-0003-3455-0919

ABSTRACT

Testing and examining go on in higher education all the time through continuous assessments and end semester examinations. The grades scored by students determine not only academic mobility but eventually who get employed in the job market, which seems to be shrinking all over the world. Those charged with testing are often staff who have higher qualifications in their subject areas but are not necessarily teaching or examination experts. Against this background, the researcher wanted to find out what was happening at selected university across three schools: Social Studies, Education and Science. The university is fairly young having been awarded its charter twenty years ago. The paper asked two questions namely, at what levels of Bloom's Taxonomy are lecturers asking examination questions? Secondly, do the level and balance of questions show growth in examining skills? The study evaluated over 1039 questions from randomly selected examination papers from the Examinations Office for the academic years from 2014/15 to 2017/18 (three academic years). A guide from the list of verbs used in Anderson s (revision of Bloom was used to analyze the questions. Descriptive statistics were used to describe the trends in testing for each year. ANOVA and t-tests were used to find out if there were significant differences between numbers across categories and within categories. The results of the study show that most examination questions are at the levels of remember (literal) and knowledge (understand). In 2016/17 and 2017/18 academic years, there were significant differences in the percentage of questions examined in these two categories. However, it seems from the study, that testing or examining skills do not grow through the practice of setting questions. There is need for examiners to be trained to use the knowledge in setting questions that discriminate effectively across the academic abilities of students they teach.

INTRODUCTION

Testing and examining are pertinent to all institutions of learning. What is taught and how effectively it is tested, however, are issues that institutions need to address. Across the world, there is still a lot of premium laid on certificates that invariably come from testing situations. The study aimed to find out how effectively testing is done at the selected university across three schools. The study sought, firstly, to examine at what levels students were being examined over three academic years. Secondly, it sought to know whether the practice in examining develops expertise that would enable lecturers (examiners) to test across Bloom's Taxonomy to bring differentiation in performance and grading of students.

Mawa, Haque and Ali (2019) argue that although the emphasis is laid on effective assessment of learning in higher education, those charged with assessing know little about what constitutes effective learning or assessment practices in higher education; the two parameters are critical if higher education is to be meaningful and lifelong. Good examination questions should test across the cognitive levels to enable both comparison and segregation of very good students from average and struggling (Tremblay, Lalancette, and Roseveare, 2012). Weir and Roberts (1994) argue that good tests elicit adequate samples of students or examinees' performance (abilities) in any one subject.

The study had two objectives: one, to find out what types of questions university students are asked in examinations with reference to Bloom's Taxonomy (Bloom, 1956; Krathwohl, 2002). Two, the study examined whether there was growth in the balance of questions across the cognitive categories as lecturers in the university gained more experience with setting examination questions.

LITERATURE REVIEW

Mawa, Haque, and Ali (2019) state that effective testing is beneficial both to students and lecturers as it gives both an assessment of the kinds of abilities that are in each class and provides feedback on curriculum coverage and whether learners have mastered the content in the subjects. This is important for future planning. It is, additionally, important at university to assess the quality of students that move from one level to the next, especially if the examinations are standardized. Assessment is considered one of the standard practices that enable lecturers and institutions to determine learning outcomes (Abosalem, 2016).

The 21st century requires quality teaching and learning especially in higher education. While teaching is mainly concerned with planning, delivery, and testing or examination of learning, quality education demands for a strong relation between subject-knowledge and skills to enable the application of the several skills using the knowledge gained. In this regard, planning, effective delivery, and assessment become critical. hence the Higher education assessment becomes

critical as institutions become production lines for resources for industry and national development agenda (Tremblay, Lalancette, and Roseveare, 2012).

Despite the importance and regularity of testing, many institutions of higher learning do not train faculty in the skills and knowledge required for effective assessment of students. The staff know little about the effective assessment of learning or assessment practices in higher education, which are crucial to a meaningful education. Training of faculty for planning, delivery and testing are often taken for granted. It is by and large assumed that if you have a PhD or Masters qualification, you qualify to teach and that you will probably teach and examine effectively as you 'learn the ropes.' An effective test or examination should meet the following criteria (Weir &Roberts, 1994; Phelps, 2006; Sweller, 1988):

i. Test across the knowledge levels of Bloom's taxonomy: knowledge, comprehension, application, analysis, synthesis, and evaluation.

ii. Solicit an adequate sample of each learner's performance in a given subject

iii. Give adequate time to each candidate to reveal what they have understood from a course

iv. Be couched in as accessible language as possible so that the language of questions does not contribute to test difficulty.

Lecturers want to measure their learners' 'knowledge' of subject content as well as their skills to relate that knowledge to multifaceted areas, tasks, skills and situations. To do this, the use of

Bloom's Taxonomy comes in handy. Using it as a guide to testing a broad spectrum of knowledge and skills in higher learning would be an advantage both for comparison across candidates, but also for gauging content mastery after teaching. To test across the categories helps each lecturer measure levels of understanding and ability to manipulate knowledge gained (Weir & Roberts, 1994; Airasian, 1994; Anderson et al., 2001).

METHODOLOGY

The current study used a cross-sectional survey design where the researcher aimed to find out how examination questions were spread across the taxonomy from three different schools: Education, Science, and Arts in three academic years and six semesters: 2014/15, 2016/17 and 2017/18 in the university. The study randomly collected past examination papers from the examination department at the university. The question past paper was analyzed using a matrix of verbs used for each cognitive category in Bloom's Taxonomy (Bloom, 1984; Anderson et al., 2001). For example, under knowledge, which generally tests simple recall of facts from content that has been taught or expected to have been gained in the course, examiners would use examples of the verbs in the following list: define, repeat, record, list, recall, name, and underline, among others. Under comprehension of content, we may have the following verbs in questions set: explain and discuss. The study used a list of verbs given in Krathwohl (2002) to isolate questions by category. A total of 1,039 questions from past examination papers were collected from examination files in the three schools as summarized below in *Table 1*.

Table 1: Samples of questions by school and year

Schools	Academic Years			Totals
	2014/2015	2016/2017	2017/2018	
Arts	107	100	97	304
Education	105	100	113	318
Science	123	93	101	317
Totals	**315**	**293**	**311**	**1,039**

From the above Table 1, it can be observed that a total of 315 questions were analyzed for the academic year 2014/15; for the academic year

2016/17, a total of 293 were analyzed, while for 2017/18, a total of 311 questions across the three schools were analyzed.

The questions were then analyzed by cognitive categories using guiding verbs from the six different levels to answer the research questions above guided by Anderson et al. (2001) and Krathwohl's (2002)) revisions of Bloom. Descriptive statistics were used to analyze the information as percentages. The data is displayed in tables and graphs. Further, an ANOVA test was run to examine whether there were any significant differences in the number of questions by year across categories and within categories.

RESULTS

Figure 1: Analysis of questions asked by category by the schools in 2014/15
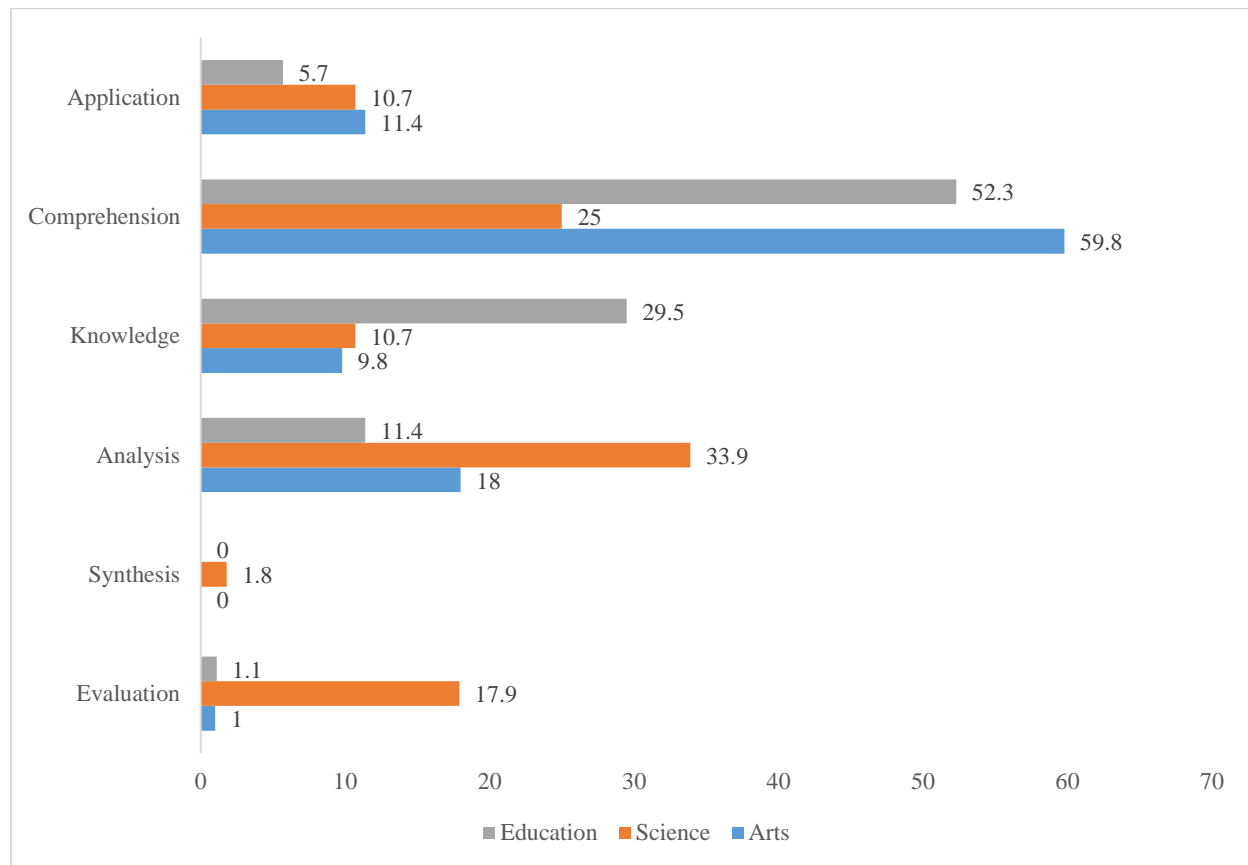
The results shown in *Figure 2* above show that in 2014/15 examinations, the school of Education asked most (52.3%) of their questions at comprehension level, the school of Arts 59.8%, and Science 25 %. The school of Education also had the largest number of questions at the knowledge level (29.5%), while Science and Arts had 10.7% and 9% in this category. The school of Science and Arts had more questions than Education at 33.9% and 18% respectively against 11.4%| at the analysis level. There were also more evaluation questions from the school of science than both Arts and Education combined for the academic year in question (see *Figure 1*).

Figure 2: Questions asked by category by the schools in 2016/17
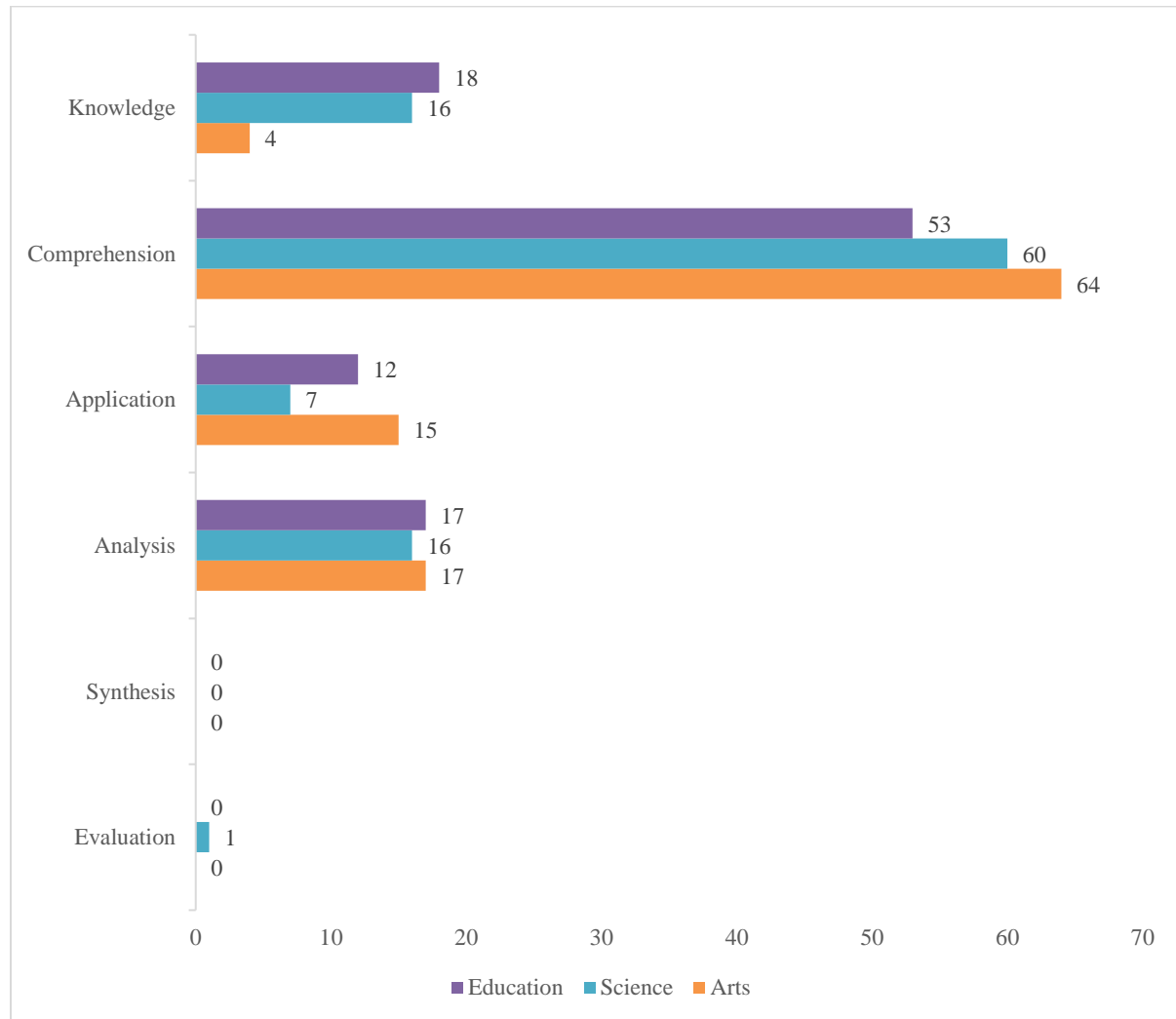
In the academic year 2016/17 examinations, *Figure 2* displays the trends in how examination questions were distributed by cognitive categories. It will be observed that knowledge questions went down significantly especially in the schools of Education and Arts. Instead, comprehension questions rose in number across the three schools at 53% for Education, 64% for arts and 60% in Science. Analysis questions went down in numbers in the school of Science from 33% to 16% but rose in the school of Education from 16% to 17% (*Figure 1* & *Figure 2*). In the school of Arts, analysis questions stayed almost at the same level, 17% and 18 % in 2014/15 and 2016/17 respectively.

The percentage of synthesis questions were negligible even in the school of Science examinations. Both schools of Arts and Education did not have synthesis questions in both 2014/15 and 2016/17 academic years. In fact, from Figure 2 above, there were almost no evaluation questions from the schools of Arts and Education. In the school of Science, evaluation questions had reduced in the two academic years from 18% to 1%, a trend that raises questions about consciousness in the need for equity or balance in the cognitive categories.

In the exam sample questions for the year 2017/18 Figure 3 below shows the distribution of questions by number and category.

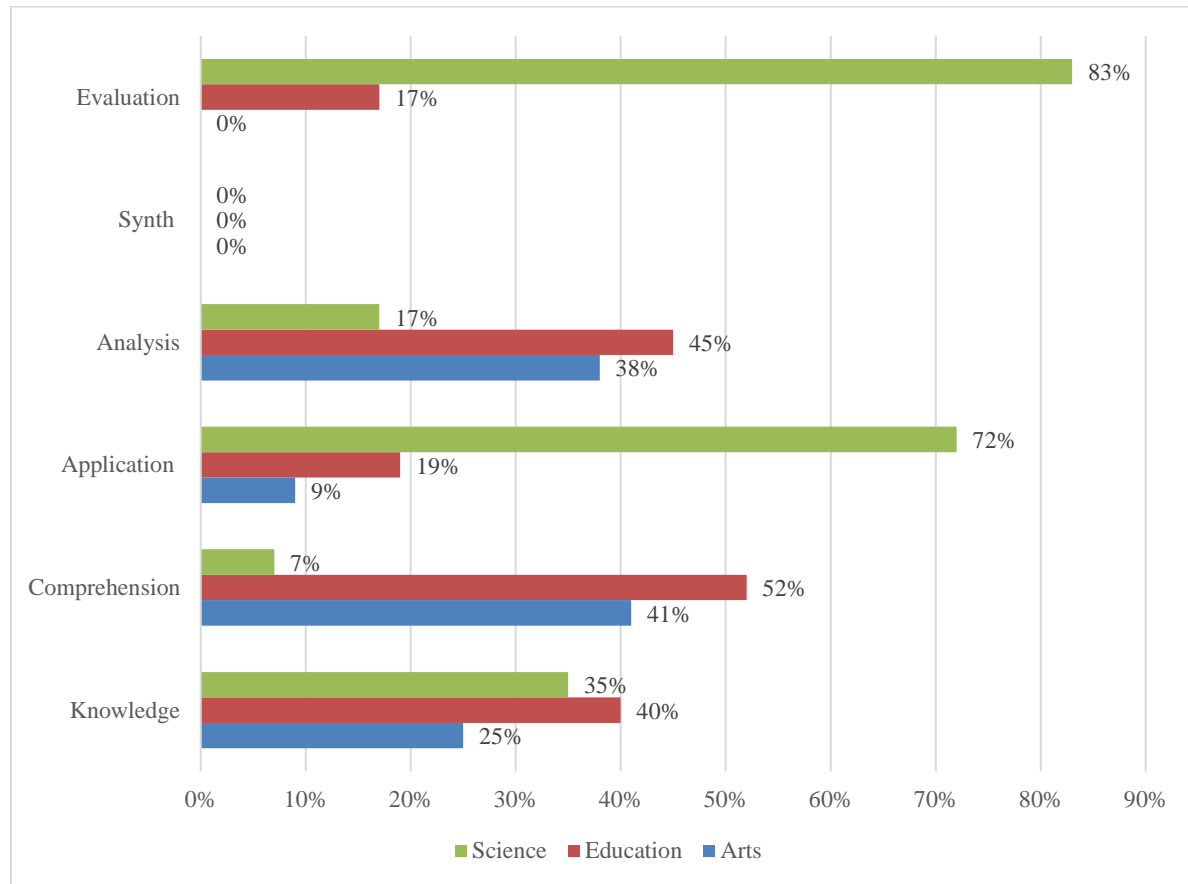**Figure 3: 2017/18 questions by school, cognitive category and percentage**

Figure 3 above summarizes the questions by category for the sample of examinations given in the academic year 2017/18. It is worth noting that in both semesters there were no questions requiring candidates to synthesize information/ knowledge in all the three sample schools- Education, Arts, and Science. Compared to the year 2016/17 there were significant increases in the number of questions testing surface-level -knowledge.

Most of the questions in the science subjects fell between knowledge and application; in fact,

application questions accounted for over 70% of the tasks the candidates had to respond to. This was over a 60% increase from the year 2016/17. In the School of Science, only 7% of the questions tested comprehension of the content that candidates had learned. While there were no evaluation questions in the School of Arts, the school of Science gave 83% of all the evaluation questions in the three schools, with the School of Education accounting for 17% from the sample questions analyzed. These trends of testing concur with the findings of Mawa, Haque, and Ali, (2019) in Bangladesh.

Table 2: Question percentages by school and cognitive level for 2016/17 and 2017/18

School	Knowledge		Comprehension		Application		Analysis		Synth	Evaluation	
	16/17	17/18	16/17	17/18	16/17	17/18	16/17	17/18	No	16/17	17/18
Arts	04	25	64	41	07	09	16	38	0	07	0
Educ	18	40	53	52	12	19	17	45	0	16	17
Science	16	35	60	07	15	72	17	17	0	16	83

Using ANOVA, the researcher tested to determine whether there were significant differences in the question numbers by category from year to year for 2016/17 and 2017/18 academic years. There were no significant differences in the numbers of different question categories in the years 2016/17 and 2017/18. However, there were significant differences only among knowledge and comprehension questions in the years 2016/17 and 2017/18 with statistics showing p 0.044, F value of 5.34 against a critical value of 4.96 at the 95% confidence level both within groups and between groups. There were, however, no differences in the numbers of questions in the other cognitive categories in the two years in question. The t-test done for the schools in the same categories found significant differences in the evaluation question category in the school of science for 2016/17 and 2017/18, the percentage of questions having risen from 16% to 83% of all evaluation questions in the year 2017/18.

DISCUSSION AND CONCLUSION

The number of questions asked per year at the university seemed adequate for eliciting reasonable samples of student knowledge and understanding. However, over 60% of the questions asked stayed at the lower levels of cognition and did not test students' abilities to re-organize information, use it in different ways, synthesize it or even apply it to novel situations (Momsen et al., 2013). The origins of this may be the following many lecturers at the university are not trained in both teaching and examination. This leaves many of them teaching and testing at the 'regurgitation' level without the questions demanding that learners do more than recall information. Some lecturers are observed dictating notes to undergraduates which often means that students will want to give back to the lecturer word-for-word what they have taken as notes during lectures in examinations. This trend does not train learners in metacognitive skills or even in critical thinking.

The absence of higher-order questions also indicates that examinations do not adequately discriminate among students' academic abilities to make it possible to say that those who get 'A's are brilliant students. It has been observed that where there are only one (01) analysis or evaluation questions among six or five lower-order questions, candidates tend to avoid the higher-order questions in favour of the lower-order questions. After all, examination scores matter enormously and passing becomes what is important.

The fact that most of the questions were only at 'remember' and 'understand at the surface' (comprehension) levels implies that students were not being tasked to reorganize knowledge, question it in radical ways, or apply it to new situations. Furthermore, this seemed to point to the fact that examining trends may reflect teaching trends that do not train students in critical thinking, metacognitive sensitivity and synthesis of factual knowledge (Momsen et al., 2013).

Similarly, it does not seem as if, from this study, awareness or ability to set 'balanced' questions naturally grows with years of teaching and examining experience in lecturers. It seems to the researcher that awareness creation for the lecturers on testing, its functions, how set 'good' questions, as well as practice in this, are important to help those who come into teaching at the university. It seems from the observation of the sample papers examined in this study, that training in testing and examining should be an ongoing activity at university. This is borne out of the fact that even lecturers of Education, who were trained in testing, seem to have forgotten the tenets that make examinations discriminatory and effective as measures of learning.

The university should train students in critical thought of the knowledge and the facts they are exposed to; furthermore, learners should be able to question new knowledge in the light of experience and reality. This should be an exciting journey for each student and lecturer -the exploration of the nuances of knowledge and facts against different realities and opposing facts. In other words, each course should expose learners to different ways of understanding knowledge presented to them, different ways of arguing with the same knowledge cogently and logically.

It is recommended that two things of necessity should be implemented: one, that all lecturers are trained in effective teaching that helps students to go behind the rudimentary aspects of understanding facts. Secondly, they should also be trained in

effective examination. Effective teaching would involve methods of teaching, ways of engaging students, even in large classes, in activities that add value to the content, that is taught through group projects and assignments. Continuous assessments that are individualized or done in groups enable students to rationalize what they are doing and why and explain the same to lecturers or fellow students. Quick activities that involve students giving solutions to problems in pairs or small groups would enable lecturers to train students in critical problem solving and help students working in pairs, for example, to rationalize solutions.

Training in effective examining would enable lecturers to be more aware of how testing relates to teaching and vice-versa. Additionally, it would also enable them to develop knowledge on how to teach and test across the cognitive categories. Effective testing only grows when one has knowledge of test construction and this is essential for each lecturer. Furthermore, examination performance provides useful feedback to the lecturer on which topics have been understood and which one's learners have difficulty learning. This information can lead to changes in approaches. However, only lecturers who have knowledge about the uses of tests as useful feedback and what to do with that feedback - will be able to use examination results to improve the 'how' of their trade. This requires something beyond awareness creation, actual training.

REFERENCES

Abosalem, Y. (2016). Assessment techniques and students' higher-order thinking skills. *International Journal of Secondary Education*, *4*(1), 1-11.

Airasian, P. W. (1994). *Classroom assessment.* New York, NY: McGraw Hill.

Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., Raths, J., & Wittrock, M.C. (Eds). (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's Taxonomy of Educational Objectives*. New York: Longman.

Bloom, B. S. (1956). *Taxonomy of Educational Objectives Handbook: The Cognitive Domain*. New York: David McKay.

Bloom, B. S. (1984). *Taxonomy of educational objectives*. Boston: Allyn and Bacon.

Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory into Practice*, *41*(4), 212-218.

Mawa, B., Haque, M. M., & Ali, M. M. (2019). Level of Learning Assessed through Written Examinations in Social Science Courses in Tertiary Education: A Study from Bangladesh. *Journal of Teacher Education and Research*, 14(1), 7-12.

Momsen, J., Offerdahl, E., Kryjevskaia, M., Montplaisir, L., Anderson, E., & Grosz, N. (2013). Using assessments to investigate and compare the nature of learning in undergraduate science courses. *CBE—Life Sciences Education*, *12*(2), 239-249.

Phelps, R. P. (2006). Characteristics of an effective student testing system. *Educational Horizons*, *85*(1), 19-29.

Sweller, J. (1988). Cognitive load during problem-solving: Effects on learning. *Cognitive science*, *12*(2), 257-285.

Tremblay, K., Lalancette, D. & Roseveare, D. (2012). *Assessment of Higher Education Learning Outcomes*. Feasibility Study Report. OECD.

Weir, C. & Roberts, J. (1994). *Evaluation in ELT*. Blackwell.