Level of Learning Assessed through Written Examinations in Social Science Courses in Tertiary Education: A Study from Bangladesh

Bentul Mawa[1], M. Mojammel Haque[2], M. Mozahar Ali[3]

[1]Associate Professor, [2,3]Professor

[1-3]Graduate Training Institute, Bangladesh Agricultural University, Mymensingh, Bangladesh

ABSTRACT

This paper examines students' cognitive learning outcomes assessed through semester final written examinations in Social Science Courses in tertiary level education. The study used a content analysis method to analyze 125-semester final written exam papers (tests) of 52 courses of B.Sc. Ag Econ. (Hons) degree program of Bangladesh Agricultural University. The study revealed that written exam papers mostly cover 'remember' and 'understanding' (18% and 60%) level of learning while 'apply-analyze-evaluate and create' levels cover only 22 percent. Year-wise change in lower order learning assessed (remember) showed a slightly decreasing trend while others showed an increasing trend to adjust that change. Level-wise (L1 to L4) increasing trend in order was observed only for 'understanding' while all others showed no definite change pattern. The study concludes that the assessment occurs mainly at lower order learning, and it does not progress with the level of studies (L1 to L4). The existing written exam strategy is not suitable to assess higher order learning to satisfy 'critical thinking and decision making' outcome so that students become better equipped for the existing job market and the rapid changing world. The program requires changing its assessment strategy to ensure higher order learning.

Keywords: Assessment, Bangladesh Agricultural University, Bloom's Taxonomy, Learning outcomes, Level of learning, Tertiary education.

Journal of Teacher Education and Research (2019). DOI: 10.36268/JTER/1413

INTRODUCTION

Modern society demands high-quality teaching and learning. Teaching-learning is mainly comprised of three major activities- teaching planning, teaching delivery, and assessment of learning. To meet standards of quality education, in addition to subject-related knowledge and skills, teachers also require a great deal of knowledge and skills with regard to effective teaching planning, delivery, and assessment. Emphasis on assessment of learning in higher education has become more important in recent years (Tremblay, Lalancette, and Roseveare, 2012). Interestingly, we know very little about the assessment of learning or assessment practices in higher education, which is crucial to a meaningful education. This paper explores this area in order to shed light on the assessment of learning in tertiary education in the context of Bangladesh.

Assessment is considered one of the standard practices to identify the learning outcomes of students (Abosalem, 2016). In general sense, a learning assessment is a method to measure how much of the knowledge and skills a learner has acquired during a course of study. Assessment of learning refers to strategies designed to confirm what students know, demonstrate whether or not they have met curriculum outcomes or the goals of their individualized programs, or to certify proficiency and make decisions about students' future programs or placements (Earl & Katz, 2006). It is designed to provide evidence of achievement to parents, other educators, the students themselves, and sometimes to outside groups, e.g., employers, other educational institutions.

Thus, the purpose of assessment of learning is to measure, certify, and report the level of students' learning, so that reasonable decisions can be made about students. Assessment of learning also provides the collection and interpretation of information about

students' accomplishments in important curricular areas, in ways that represent the nature and complexity of the intended learning. In other words, genuine learning for understanding is much more than just recognition or recall of facts or algorithms. Assessment of learning tasks needs to enable students to show the complexity of their understanding. Students should able to apply key concepts, knowledge, skills, and attitudes in ways that are authentic and consistent with current thinking in the knowledge domain (Earl & Katz, 2006).

On the whole, assessment is not only considered as an important tool for reporting a particular student's performance but also it helps to evaluate the performance of the entire system. Authors, like Airasian (1994), Pellegrino, Chudowsky & Glaser (2001) note that assessment has three main purposes: to assist learning, to measure a particular student's achievement and to evaluate the whole program. Without good assessment techniques, it is difficult to ascertain whether reforms in teaching-learning or curriculum are working.

In addition, assessment of learning is an ongoing process as it is being conducted continually in various forms. Methods for assessment of learning may include not only tests and examinations,

but also include a rich variety of products and demonstrations of learning, such as portfolios, exhibitions, performances, presentations, simulations, multimedia projects, and a variety of other written, oral, and visual methods (Earl & Katz, 2006). However, at the University level in Bangladesh, written assessment strategies are considered the dominant one. Written assessment strategies are practiced in different points of learning timeline, such as class-test, midterms, quizzes, and semester final examinations. Among these written methods, semester final assessment is the major form of assessment as it includes 70% of weight in total marks distributed.

Critics argue that the written tests are poor because through this test only the verbal ability of students is tested (Wiggins, 1994). This test is usually a one-time measure and is based on the achievement made by a given student on a particular day. This type of assessments usually relies on a student's single correct answer per specific question (Wraga, 1994), usually omitting the student's demonstration of overall knowledge and their thought process. Therefore, in recent years there has been a demand for better methods of assessing students achievements in order to measure what students can do with what they know, rather than simply finding out what they know (Struyven, Dochy, Janssens, Schelfhout, & Gielen, 2006; Boyd, 2008). As a result of this demand, authentic or performance-based assessment has emerged, which comprises the assessment of traditional academic content in combination with the skills and knowledge essential for lifelong learning. These new forms require students to construct responses orally or in writing to a wide range of problems, create a product, or demonstrate the application of knowledge in an authentic context (Calfee, 1994).

In other words, it can be said that authentic forms of assessment encourage students to use higher order cognitive skills to use their knowledge creatively. Educators consider higher order cognitive skills as high order thinking that occurs when the student obtains new knowledge and stores it in his memory, then this knowledge is correlated, organized, or evaluated to achieve a specific purpose. These skills have to include sub-skills such as analysis, synthesis, and evaluation, which are the highest orders in Bloom's (1956) cognitive taxonomy. Benjamin (2008) argues that these skills can be developed in a cumulative fashion as students progress through their courses and subjects and other experiences they get from their institutions. As well, by including their subjects by problem-solving, critical thinking, and decision making activities help students enhance their higher-order thinking skills. The term 'higher order' thinking skills suggests that there is another set of 'lower order' skills that need to come first. Newman (1990) differentiates between the two categories of skills concluding that the lower skills require simple applications and routine steps. For example, it includes remember, understanding, and application. In contrast, higher order thinking skills challenge students to interpret, analyze, or manipulate information (Newman, 1990).

There is no doubt that the development of student's higher-order thinking is the central goal for all educators and educational stakeholders. In order to uplift the standard of our higher education, heightened cognitive skills are very important at the tertiary level. The Self Assessment Manual (2016) of University Grants Commission (UGC)—Bangladesh determined the national requirements outlining that teaching-learning in higher education must address the higher order of learning in the educational domain, i.e., 'application to creation' and efforts should be taken to develop the skills of original thinking and creative faculty.

While assessment of learning appears to be among the most important pieces of information on higher education, available data remain scarce in the literature (Tremblay et al., 2012). Very few studies have been conducted worldwide to examine the assessment of learning at the college level (Stiggins, Griswold, & Wikelund, 1989; Boyd, 2008; Abosalem, 2016). To date, there is no research evidence at University level, which examined whether semester final written examinations address higher order of learning or not and therefore cannot inform us about this. Thus, this research aims to examine the order of cognitive learning outcomes assessed through semester final written examinations in Social Science Courses of Bangladesh Agricultural University (BAU). The study has been designed with the following specific objectives:

• Determine the extent of learning assessed under a different order of cognitive domain through semester final written examinations.
• Verify if there is a trend of improvement in cognitive learning assessment under successive years.
• Verify if level-wise (L1 to L4) increasing trend in higher order learning (and a decreasing trend in lower order learning) exists in semester final written examinations.

To meet the globalization challenges raising higher education quality to the world standard is essential. Bangladesh Government plans to prepare university graduates in such a way that they can successfully compete in the context of the international knowledge society. Through this study, we will be able to know our existing situation with regard to assessment practices at the university level. This will further aid educators and policymakers to take necessary corrective measures in developing standard assessment plans for future students of tertiary education.

METHODOLOGY

This study was carried out at the Graduate Training Institute (GTI) of BAU, Mymensingh during 2017–2018. In order to examine student's extent of learning under different order of cognitive domain, or to know how teachers construct questions, how much of them agreed with the order of Bloom's Taxonomy, whether there are any year-wise and level-wise differences in Bloom's cognition levels, all the semester final questions on 52 courses under the B.Sc. Ag Econ. (Hons) degree program, together with the courses offered by the Department of Agricultural Extension Education were collected.

It is here to be noted that the Faculty of Agricultural Economics and Rural Sociology, one of the six faculties of the Bangladesh Agricultural University was established in 1963 with five academic departments: Agricultural Economics, Agricultural Finance, Agricultural Statistics, Agribusiness and Marketing, and Rural Sociology. The faculty offers a four-year undergraduate B.Sc. Ag. Econ. (Hons.) degree program. These five academic departments offer 45 courses (including eight collateral courses) (Table 1). Questions from all the core courses were chosen for this study. Among eight collateral courses three courses, i.e., English Language, Computer Application in Social Science and Agricultural Extension were also selected; however, other five collateral courses relating to Agricultural Science were excluded for the purpose of this study. In addition, questions from 07 elective courses, i.e., Legal Environment of Business, Agricultural Finance II, Consumer Behaviour, Social Structure of Bangladesh, Agribusiness Management, Local Level Planning, and Evaluation and Micro Credit offered for this program to Level-4 (Year) students were examined in this study. Also, questions from five courses, such as Fundamentals of Extension, Leadership and Motivation, Extension Communication and Group Approaches, Agricultural Extension Education, and Extension Organization Management were selected from the Department of Agricultural Extension Education, who offers these courses to

Table 1: Number of departments and courses offered for B.Sc. Ag Econ. (Hons) program

Departments	Courses
1. Agricultural Economics	Principles of Economics, State and Government, Public Administration and Political History of Bangladesh, Microeconomics I, Microeconomics II, Macroeconomics I, Mathematical Economics, Macroeconomics II, Production Economics I, Farm Management I, Economy of Bangladesh, Economic Growth and Development, Environmental Economics, Farm Management II, Production Economics II.
2. Agricultural Finance	Money and Banking, Public Finance, Agricultural Finance I, International Economics, Agricultural Policy and Planning, Land Resource Economics, Financial Management.
3. Agricultural Statistics	Descriptive Statistics, Inferential Statistics, Regression Analysis, Economic Statistics/ Econometrics.
4. Agribusiness and Marketing	Management Accounting, Principles of Marketing, Business Principles and Entrepreneurship Development, Cooperatives in Theory and Practice, Agricultural Price Analysis, Institutional Economics and Rural Development, Agricultural Marketing, Introduction to Agribusiness, Research Methods in Agricultural Economics.
5. Rural Sociology	Principles of Sociology, Rural Sociology.

Source: B.Sc. Ag. Econ. (Hons.) degree program's Self-Assessment Report 2018

the faculty of Agricultural Economics and Rural Sociology and other faculties of Bangladesh Agricultural University. In total, 125 questions were used for analysis.

The questions were collected from BAU library. Table 2 shows the number of level-wise questions collected for the study.

The questions were analyzed adopting content analysis. Content analysis is a widely used method to analyze text data. Research using content analysis focuses on the characteristics of language as communication with attention to the content or contextual meaning of the text (Budd, Thorp, & Donohew, 1967; McTavish & Pirro, 1990; Tesch, 1990). There are three distinct approaches to content analysis: conventional, directed, and

Table 2: No. of questions undertaken level-wise in 3 years (2014–2016)

Level-wise questions	No. of questions
Level 1	19
Level 2	27
Level 3	35
Level 4	44
Total	125

summative. All three approaches are used to interpret meaning from the content of text data. In conventional content analysis, coding categories are derived directly from the text data. With a directed approach, the analysis starts with a theory or relevant research findings as guidance for initial codes. Summative content analysis involves counting and comparisons, usually of keywords or context, followed by the interpretation of the underlying context (Hsieh & Shannon, 2005).

In this research, researchers chose a directed approach to content analysis. This approach is appropriate when existing theory exists about a phenomenon and helps to determine the initial coding scheme. Bloom's (1956) six level of cognitive domain served as an initial framework in this study to examine the extent of both the lower order (knowledge, comprehension, and application) and higher order (analysis, evaluation and synthesis) learning that students have to gain and acquire at the end of their courses. Constructing tests according to this behavioral approach requires teachers to show their competence in each content area.

After collecting all the questions, coding started immediately by reading each question several times carefully. The next step was to highlight action verbs that appeared to describe any six orders of the cognitive domain. All highlighted text was coded using the predetermined categories wherever possible. For example, there were questions like, what are the differences between state and government? Or how would you define sovereignty? The researcher very carefully began to label six levels/orders of cognitive processes within these questions. However, the difficulties researchers faced here that in many cases actions verbs were not enough to determine the cognitive level; thus the actual meaning of the questions determined the order of cognitive domain. Also, data that could not be coded into one of six categories derived from the theory was reexamined and coded with another label that captured the problems found in the questions. After coding all the questions data were recorded in Microsoft Excel for analysis. Finally, researchers compared the extent to which the data were supportive or contradictory to other research available and what new perspectives were added.

RESULTS AND DISCUSSION

Standard 5–9 of SA Manual (2016) of UGC states that assessment procedure should be comprised of a set of multiple activities to measure the attainment of students' learning outcomes and skills. In practice, teachers usually undertake a limited number of assessment strategies, such as class test and final written examination for assessing student's cognitive learning. Marks distributions are as follows: out of 100, 20 marks are allocated for class test. In some cases, some teachers distribute the marks of the class test into the assignment and presentations. Ten (10) marks are kept for attendance, and the major portion marks (70%) are allocated for final written examinations. One of the major aims to semester final written examination is to assess learning outcomes of students in relation to contents covered in the course rather than developing their potential in terms of employability. The assessment procedure must be designed to test the abilities and skills of student for integration and application of knowledge and analytical approaches (Standard 5-10: SA Manual, 2016).

This study explored whether semester final written examination at the university level in Bangladesh agreed with Bloom's six major levels of cognitive order or not (Graph 1). Result of this research suggested that 18% test items were 'remember' items, 60% belonged to 'understand' level questions, and 12% of them
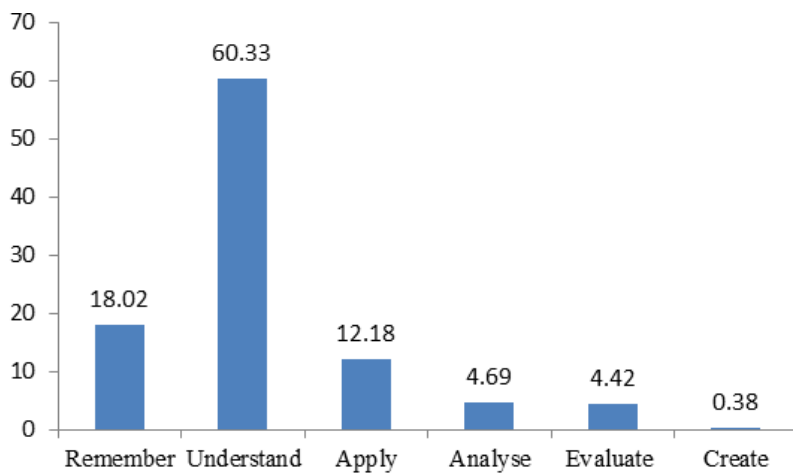
Graph 1: Percentages of marks distributed under different levels of the cognitive domain

were 'application' level items and these three levels cover 90% of the total. Only 10% items (analyze—4.69%, evaluate—4.42%, and create—0.38%) fell into higher-order thinking skills, which indicated that the majority of teachers' assessment items focused on the lower three levels of Bloom's Taxonomy. These findings reflect the tendency of the assessment strategy used in social science courses of BAU to ask students to recall information or to explain a topic, which would not help students in improving their higher-order thinking skills.

Findings of this study agree with Stiggins et al., (1989) research. Stiggins et al., (1989) studied the assessment practices (use of oral and test questions) of 36 teachers who taught mathematics, science, social studies, and language arts at grade levels 1 to 12. Excluding math items, over one-half of the test questions at all grade levels were recalled measures (55%) followed by inference (19%), analysis (16%), comparison (5%), and evaluation (5%). In math, 72% of the questions measured inference, 19% measured recall, and 9% measured comparison. Oral questions followed the same pattern, with slightly less than half of them measuring recall. They found that few items measure skills above the third level of the taxonomy, and a high percentage of them recall (knowledge) items, thereby concluding that teacher-made tests are not generally of high quality and that teacher assessment of higher order thinking skills is rare.

Another aim of this research was to identify the year-wise differences in the level of questions as per Bloom. Findings of this research suggested that there was no significant difference between the years from 2014 and 2016 (Table 3). There was a minimal change in remember level items which was 19% in 2014 and it reduced to 16% in 2016. It might be the fact that in recent years UGC is giving emphasis on quality assurance in higher education. It also realizes that the state has the responsibility to put in place an enabling framework that would encourage tertiary education institutions to be more innovative and responsive to the demanding needs for rapid economic growth and to empower the graduates with right skills for successfully competing in the global

knowledge economy (UGC 2015). Against this backdrop the Ministry of Education (MoE), the UGC and the HEIs felt that it is high time to establish institutional quality assurance cell (IQAC) in every public and private university to develop a mechanism for the systematic review of study programs, to ensure quality teaching-learning, research, knowledge generation and support services standards at an acceptable level.

In addition, this research verified if level-wise (L1 to L4) increasing trend in higher order learning (and a decreasing trend in lower order learning) exists in semester final written examinations (Graph 2). There was no definite increasing or decreasing trend in remember level; however, there was an increasing trend in an understanding level ranging from 49.66 percent to 64.70 percent. There was no increasing or decreasing trend from applying to create a level.

The possible reason for this might be teachers are either not aware of Bloom's (both lower and higher order) cognition levels or they are not willing to prepare questions adopting Bloom's Taxonomy. This is because constructing questions according to Bloom's (1956) Taxonomy requires not only time but also requires much more creativity. Faculties should be more careful and spend much time in constructing creative questions.
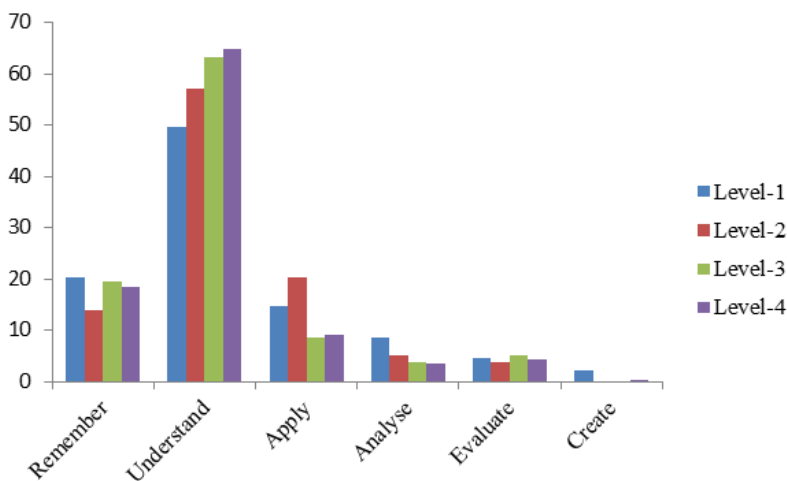
Problems in Questions

As noted earlier, researchers identified several problems while coding questions which are mentioned here. The content analysis suggested that semester final written test questions have validity and reliability problems. In terms of validity, many tests did not cover questions from six levels of Bloom. With regard to reliability, marks were not distributed for each part of the questions. Even in some cases, marks were not given on individual questions. In those cases, only total marks were mentioned in the heading of the question paper. As the student progresses through their levels and courses, the teacher should allocate fewer marks for lower order items and more marks for higher order. However, the findings of this study showed that the opposite was true.

Table 3: Year-wise distribution of marks under six different levels of the cognitive domain

Year	Remember (%)	Understand (%)	Apply (%)	Analyze (%)	Evaluate (%)	Create (%)
2014	19	60	12	4	5	0
2015	18	61	12	5	4	0
2016	16	59	13	6	5	0

Graph 2: Level-wise distribution of marks under six different levels of cognitive domain

Wrong use of action verbs in setting questions was another problem. Many questions started with the action verb 'analyze' but those were not the questions for analyzing the topic. In some cases, questions were vague; for example, describe the causes and risk factors of crime following different theories and approaches including the WHO ecological model. This erroneous approach in setting questions indicates a vague understanding of the academics on the notion of preparing creative questions.

Conclusion

Teaching-learning should be need-based, practical oriented, and involve aspects of critical thinking and inspire students to apply acquired knowledge in real-life situations. In this regard, SA Manual of UGC (2016) outlines in Standard 5–8 that student performance assessment approach must be focused on higher-order learning. However, in this study order of learning assessed through the semester, final written examinations mostly cover remember (18%) and understanding (60%) while that of apply, analyze, evaluate and create covers only 22 percent. Year-wise change in lower order learning assessed (remember) shows a slightly decreasing trend while others show an increasing trend to adjust that change. Level-wise (L1 to L4) increasing trend in order was observed only for understanding level (cognitive-2) while all others showed no definite change pattern. Thus, it is evident that written examination question papers show significant emphasis on lower order (remember and understand). Student performance assessment must be designed to test the abilities for integration, application of knowledge and analytical approaches so that they can compete in the real world and therefore, cannot test student's analytical skills towards different aspects. Assessment should be designed in a way such as arranging group discussion, enabling students to provide individual comments, involving them in project writing on different practical or field oriented problems so that they can apply their analytical skills.

Recommendations

Based on the findings, this study makes some recommendations:
• Semester final written examination is fit only for lower order learning- remember and understand. To include assessment of higher order learning, different types of continuous assessment strategies such as exercise, case studies, assignment, and projects need to be adopted frequently.

• At initial levels (levels 1 and 2) lower order learning may occupy a major portion of the assessment, however, in the higher level of studies (levels 3 and 4) higher order learning should be increased.
• To improve the assessment of learning, different strategies should be incorporated, and to that end introduction of 'Table of Specification (TOS)' should be formalized.
• Training facilities of faculties on pedagogy is rare. Faculty members should be provided with training on innovative teaching-learning and assessment approaches with an emphasis on the preparation of creative questions.

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.

Benjamin, R. (2008). The Case for Comparative Institutional Assessment of Higher-Order Thinking Skills. Change, 40(6), 51-55.

Beyer, B. (1983). Common sense about teaching thinking. Educational Leadership, 41(3), 44-49.

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

Boyd, B. (2008). Effects of state tests on classroom test items in mathematics, School Science and Mathematics, 108 (6), 251-261.

Budd, R. W., Thorp, R. K., & Donohew, L. (1967). Content analysis of communications. New York: Macmillan.

Calfee, R. C. (1994). Cognitive assessment of classroom learning. Education and Urban Society, 26(4), 340-351.

Earl, L., & Katz, S. (2006). Rethinking Classroom Assessment With Purpose in Mind. Winnipeg, Manitoba: Western Northern Canadian Protocol.

Hsieh, H. & Shannon, S. E. (2005). Three Approaches to Qualitative Content Analysis. Qualitative Health Research, 15(9), 1277-1288.

McTavish, D. G., & Pirro, E. B. (1990). Contextual content analysis. Quality and Quantity, 24, 245-265.

Newman, F. M. (1990). Higher Order Thinking in Teaching Social Studies: A rationale for the assessment of Classroom Thoughtfulness. Journal of Curriculum Studies, 22(1), 41-56.

Pellegrino, J., Chudowsky, N. & Glaser, R. (Eds.). (2001). Knowing what students know: The science and design of educational assessment: A report of the National Research Council. Washington DC: National Academy Press.

Self Assessment Manual. (2016). University Grants Commission of Bangladesh. Ministry of Education. Higher Education Quality

Enhancement Project (HEQEP). University Grants Commission of Bangladesh.

Self Assessment Report. (2018). B.Sc. Ag. Econ. (Hons.) Degree Program, Faculty of Agricultural Economics and Rural Sociology, Bangladesh Agricultural University.

Stiggins, R. J., Frisbie, D. A. & Griswold, P. A. (1989). Inside high school grading practices: Building a research agenda. Educational Measurement: Issues and Practices, 8(2), 5-14.

Struyven, K., Dochy, F., Janssens, S., Schelfhout, W., & Gielen, S. (2006). The overall effects of end-of-course assessment on student performance: A comparison between multiple-choice testing, peer assessments, case–based assessment and portfolio assessment. Studies in Educational Evaluation, 32(3), 202-222.

Tesch, R. (1990). Qualitative research: Analysis types and software tools. Bristol: PA: Falmer.

Tremblay, K., Lalancette, D. & Roseveare, D. (2012). Assessment of Higher Education Learning Outcomes. Feasibility Study Report, OECD. Retrieved from http://www.oecd.org/education/skills-beyond-school/AHELOFSReportVolume1.pdf

U.G.C. (2015). Background of the Quality Assurance Unit (QAU). University Grants Commission of Bangladesh. Retrieved from http://www.qau.gov.bd/page/background-quality-assurance-unit-qau\

Wiggins, G. (1994). Toward more authentic assessment of language performances. In C. Hancock (Ed.), Teaching, testing, and assessment: Making the connection. Northeast Conference Reports. Lincolnwood, IL: National Textbook Co.

Wraga, W. G. (1994). Performance assessment: A golden opportunity to improve the future. NASSP Bulletin, 78(563), 71-79.