The Implementation of "Test of Evaluating" and "Test of Creating" in the Assessment of Learning by EFL Lecturers in Pandemic Era

Yulis setyowati[1], Dwi Fita Heriyawati[2], Deny Kuswahono[3]
yulissetyowati@uwp.ac.id
[1,3] Language and Literature, Wijaya Putra University, Surabaya, Indonesia
[2] Faculty of Teacher Training and Education, University of Islam Malang, Malang, Indonesia

Abstract

This research aims at analyzing the implementation of a test of evaluating and test of creating as contained in the mid-term test and final test exercises during the era of Covid-19 pandemic by two English Foreign Language Lecturers. This research uses descriptive qualitative method with document analysis approach intended in analyzing the implementation of a test of evaluating and test of creating as contained in the mid-term test and final test exercises. The source data of this study is the mid-term tests and final tests that were made during the era of Covid-19 pandemic by two English Foreign Language Lecturers who teach in the English Literature department in a private University in Java. Data collecting technique is done using documentation and observation methods. The data analysis has been done by data coding. Then, the data is identified based on the categorization of level competency based on the Bloom taxonomy, particularly the evaluating and creating competency level. After that, the data will be analyzed and interpreted. The finding showed that 'test of evaluating' appeared only one item that requires students to review especially criticizing based on the qualitative standard and criteria. Meanwhile, the second finding was 'test of creating' in mid-term and final test patterned on the ability to produce a new thing by implying the processes inside 'creating' that are 'generating, planning, and producing'. By implicating Bloom taxonomy, the lecturers are not only assisted but also expanded and sharpened the assessment, so the purpose of learning is not just teaching transfer but also covering all the cognitive processes needed in learning transfer.

Introduction

The teaching phenomenon in the 21ˢᵗ century, teachers or lecturers should prioritize the quality standard such as content, process and assessment to produce an outcome as the society's expectation. As for assessment, should focus on measurement, instrument along with assessment and evaluation methods. Assessment and teaching cannot be separated, they are integrated into one system. Teaching quality can be seen in assessment quality. For this reason, an assessment should be designed in order to fit to 21ˢᵗ century era.

In the real world graduates must have soft skills as well as hard skills. This statement is supported by Ledward and Hirata's research (2011) stated that teaching achievement in 21ˢᵗ century is students should be able to live and have career. The abilities to live and have career lies on having at skills to synthesize information, work in a team, manage widely and complex (Lindemann, D.J &Tepper, J. S. 2012), and responsible for the community and the surroundings (Kandappan. &Nitin in Akib and Muhsin, 2019). In Indonesia, assessment aimed to determine students' level of competency, skill and performance. To achieve this goal, an appropriate and proportional assessment is needed. Assessment of learning reflects the society's needs, so it requires teachers to prepare a suitable assessment in this millennial era. In this context, assessment of learning as a factual data transmission and skill from a lecturer to the students covers the basic knowledge, knowledge process, knowledge content and problem solving.

Assessment of learning is an assessment based on product and mostly done in every end of semester and as an example which used widely is test. The tests that will be used in this study are mid-term test and final test since these kinds of assessment are used by nearly all of the universities to know the result of students' learning. Some studies found out that Bloom taxonomy is the most effective tool to develop assessment (Köksal & Ulum, 2018). Bloom taxonomy can show the confirmation of cognitive competency level that teachers are expecting in the end of teaching (Köksal & Ulum, 2018).

This study is focused only on the level of competency cognitive of evaluating and creating. These two cognitive competency levels are chosen for the close relation between evaluating and creating level and critical thinking (Setyowati, 2019). (Aghaei, 2018), (Thomas, 2017), (Zuhri, 2018) and (Ebadi, 2016) found that in particular, the skills of analyzing, evaluating and creating are the indicators of critical thinking. Thus, the usage of taxonomy Bloom is shown in the increment of students' critical thinking. Critical thinking skill only happens when students activate in analyzing, evaluating and creating (Aghaei, 2018).

Critical thinking as one vital aspect in higher education field is a must have competency for university graduates since they are demanded to be able to think

critically in the world of work. This is why critical thinking is an important skill to have. Some researches related to assessment and the revised Bloom's taxonomy have been done, one of them by Kamlasi and Sahan (2018) who found that in the arrangement of teacher's test should distribute the domains of Bloom taxonomy. Although there are many studies done on assessment and Bloom taxonomy, yet there is no research which focuses on the usage of the evaluating and creating cognitive domain that written in test item. The differences between the previous researches and this research are on the qualitative approach. This study focuses on the data test items in the midterm test and final test exercises by using taxonomy Bloom theory. Thus this study will try to describe the implementation of test of evaluating and test of creating which have been applied in midterm and final test by the EFL lecturers.

Assessment in Taxonomy Bloom

Taxonomy Bloom is a framework categorization of learning purpose created by Benjamin Bloom in 1956. It refers to the clarification of statements uses to predict students' ability of studying as the result of learning process. This taxonomy then revised by Lorin Anderson around 1990. Anderson, et.al. (2001) the revision covers remembering, understanding, applying, analyzing, evaluating and creating. The picture below shows the taxonomy in detail.
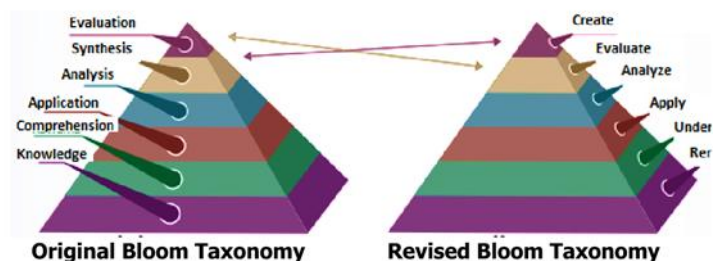
Assessment is a process of collecting data that focuses on grading. This is the final and formal way to determine object on all grading aspects of teaching process (Krathwohl, 2014). This assessment gives hope in students' success or failure in studying. The result is actionable to evaluate learning process to be better in the future.

Taxonomy Bloom gives guidance for lecturers to develop assessment which explores students' critical thinking by developing test of evaluating and test of creating. Test is an assessment tool that is widely used in educational institutes to determine the result of learning process. The ideal test is a test that accommodates the standardize competency level.

Test of evaluating

Evaluating defines as assessment based on criteria and standards. The most used criteria are quality, effectivity, efficiency and consistency. Standard can be in

quantitative (i.e. Is it a sufficient amount?) or qualitative (that is, Is it good enough?). This category covers the process of cognitive examining (refers to the assessment of internal consistency) and critic (refers to the assessment of external criterion). Examination (also called coordination, detection, monitoring or testing) happens when a student detects inconsistency or error in a process or product, determines whether a process or product has internal consistency, or detects the effectiveness of an outgoing procedure. When it combined with planning (cognitive process in category, make) and implementation (cognitive process in category, apply), checking involved determining how good the plan works. Critic (also called judging) happens when a student detects inconsistency between a product or operation and some external criteria, determines whether a product has external consistency, or values the appropriate procedure for the given problem. Criticizing lays on the core of critical thought. In criticizing, students review the advantage of a product or operation based on criterion and standard determined by students (Anderson, Lorin W and Krathwohl, 2001), (Cullinane, 2015), (Eber& Parker, 1999).

Evaluating describes as grading based on requisite and norm. Quality, affectivity, efficiency and accuracy are the most used measurements. Norm can be in quantitative form (i.e. Is it sufficient?) or qualitative (i.e. Is it good or enough?). This skill involving cognitive process to consider (refers to the decision of internal consistency and criticizing (refers to decision which relies on external requisite). The evaluation of decision making succeeds and makes assessment reflective to be inputted into critical thinking dimension (Paul, R., & Elder, L., 2008). Students are demanded to do or review product, diary or critic.

Test for evaluating refers a test which requires test takers to justify, check, judge a decision or course. The test let test taker to perform their competence to check process or product on how good they are running.   Thus the test takers also are demanded to do critiquing a product or procedure and some outer standards between a product or procedure and outer standards, of a process for an issue presented.

The item test typically used the following verb:

Appraise, argue, assess, choose, compare, conclude, criticize, critique, defend, estimate, evaluate, judge, manage, prepare, rearrange, reconcile, set up, synthesize (Stanny, 2016)

Test of Creating

Creating including placing collective component to make a whole consistency or operational; that is, reorganizing component to be model or a new framework. It involved making students to create initial item. In this creating level, activity that upgrades students' critical thinking is students should be exposed to create, build, design, produce something new as in research project activity, performance test, essay writing or product designing. Creating involved joining elements to form

coherent or functional totality; that is, rearranging element to be new pattern or structure. The classified purpose as Make, engaged students to make the original product. Composition (including writing), for instance, often, but not always, involve cognitive process related to creating. It can be in the form of procedural knowledge application (i.e. "Write this essay in this way").

Creative process can be divided into three phases: (a) problem representation, where a student tries to understand the assignment and produce an impossible solution; (b) solution planning, where a student checks the possibility and arranges plan that can be applied; and (c) solution realization, where a student succeeds to perform the plan. Thereby, creative process can be deemed to begin with divergent phase where various solutions may be reputed as student's effort to understand assignment (producing). It follows by convergent phase, where solution method designed and transformed into action plan (planning).

Plan in the creating runs when the solution is built (producing). So, creating process can be broke into three cognitive processes: resulting, planning and producing. Resulting (also called hypothesizing) involved creating alternative hypothesis based on criteria. When resulting overreaches the limit or the previous knowledge limitation and the available theory, it involves different thoughts and create the core of what is known as creative thinking. In resulting, a student is given a problem description and should make alternative solution. Planning (also called designing) engaged designing method to finish some tasks. However, planning stops doing steps to create actual solution to divide task into subtask that should be done when solving problem). Teacher often passes stating planning purpose, on the contrary stating their purpose in term of production, the last step of creative process. When it happens, planning might be done by students secretly, in order to build a product (which is, producing). In planning, a student develop solution method when the problem statement is given.(Anderson, Lorin W and Krathwohl, 2001), (Leslie, 2016), (Cullinane, 2015)

Test for creating requires test takers to generate or produce new ideas, products or ways of viewing things through constructing. Thus, in this kind of test, test takers are demanded to perform their capability in rearranging components into a new arrangement or construction. This takes students' creation or having an unique invention.
The item test typically used the following verb:

Arrange, assemble, combine, compose, construct, create, design, develop, devise, formulate, generate, invent, modify, organize, plan, prepare, produce, rate, revise, write. (Stanny, 2016)

Method

This research methods are analyzing and describing the implementation of test of evaluating and test of creating on mid-term test and final test exercises which arranged by English lecturer. Hence, this research is done and categorized as

qualitative descriptive research. There are five English lecturers whom the mid-term and final test exercises used as research object. The total of mid-term and final test exercises examined are twenty-one (21) exercises. The research is done by collecting, classifying, analyzing data and summarizing and report writing. The source of data in this research is data document in the form of mid-term and final test exercises. Prior (2003) in Cohen 2007: 201 stated that document is beneficial in presenting the phenomenon in education world. In analyzing the data, researcher applied these steps:

1) Codification – researcher gives code on each test item written on mid-term and final test exercises, 2) Classification – researcher classifies data into category which based on Taxonomy Bloom, especially which related to test of evaluating and test of creating, 3) Data analysis – researcher analyses and interprets the categorized data, 4) Discussion – researcher discusses the result finding by describing the implementation of test of evaluating and test of creating usage as well as the mapping into graphic. The research process can be seen in the table below:
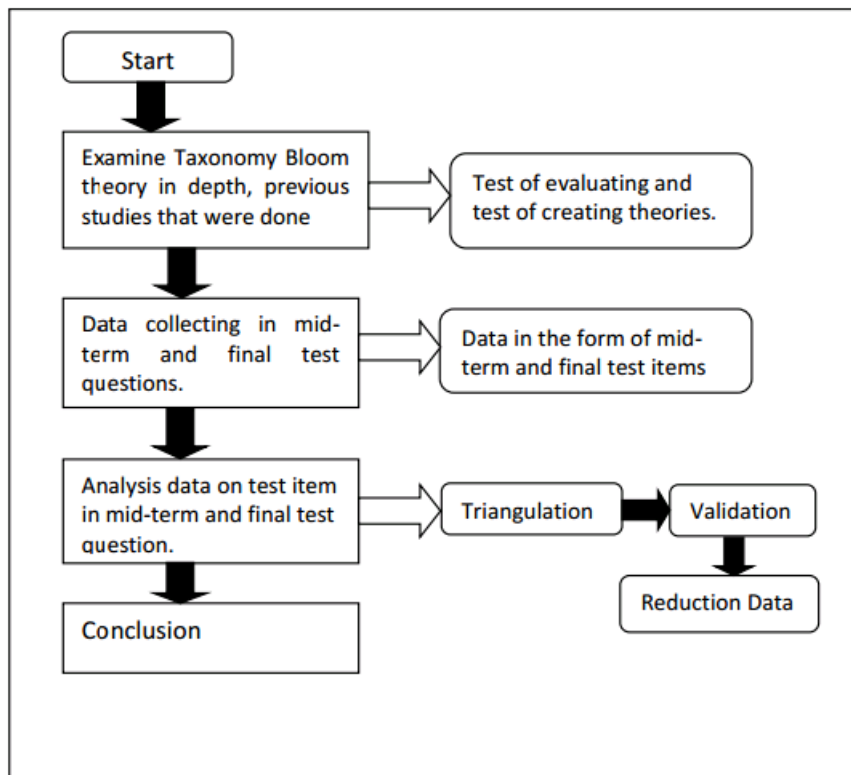
Result and Discussion

In designing the learning assessment during pandemic, lecturers are demanded to adjust the condition and current development. The result of data analysis shows that lecturers already arranged the assessment with the taxonomy Bloom scale on cognitive domain particularly on test of evaluating and test of creating. As for the usage of verb in the test of evaluating and test of creating can be seen in the table below:

Test of evaluating- (make judgment based on criteria and standards)	frequency	Test of creating (put elements together to form a coherent or functional whole, reorganize elements into a new pattern or structure)	frequency
1. Checking		1. Generating	
• Coordinating	-	• Hypothesizing	-
• Detecting	-	2. Planning	
• Monitoring	-	• Designing	-
• Testing	-	3. Producing	
2. Critiquing		• Constructing	5
• Judging	1		

The table above, shows that from the questions analyzed, is found only one item shows test of evaluating. As for test of creating there are six exercises which inquire students to show their ability in completing test of creating. One out of three English lecturers uses test of evaluating and only one exercises. While the two lecturers make test of creating. In this test of creating is found five items of exercises made by the first subject and one item of exercises made by the third subject. Test of evaluating is made to inquire students to criticize several topics. The exercise items written as follow:

Write 500 to 600 words summary and comments on the issue of translation as follow

According to the presented test items above, lecturers use test of evaluating by inquiring students to review based on criteria and standards. Students are inquired to write a script which criticize translation issue. In this term, lecturer uses the verb "critiquing". In the exercise items above, require students to be able to evaluate in term of criticizing the material discussed in the lecturing. The evaluation can be done on content, idea, concept and how to respond related issues. Evaluation from the point of view can evaluate effectiveness in criticizing the given issues. In this level of test of evaluating, a deep understanding of the discussed material of the lecturing is not the only one they need, but also their knowledge outside lecturing. In another word, test of evaluating also needs high level of

cognitive activity. It is better to be used to measure students' ways of thinking, and to train them well. Students are inquired to measure their ways of thinking, predicting, presenting and holding their views by making assessment on information, idea validity or occupation quality based on the set of criterion.

Test of creating was found six items of exercise made by two lecturers. Test of creating requires students to be able to create ideas and new information using the material that has been studied before. Students should be able to correlate or generalize the available concept, case, idea, fact and opinion, collect the joining information in different ways by combining elements in a new pattern or an alternative suggestion. The test on this level causes students to actuate in finding new information, predicting and solving problem.

Cognitive activity on this level is a high level and complex activity. Based on the exercise items presented in mid-term and final test, shows students' way of thinking. Test of creating gives a chance to students to show their creative way of thinking, creativity to combine cases, to join between concept and fact, to generalize and so on. Here, students should fully understand the material or problem presented. In this case, lecturers train their students well, not only to think critically on a problem but also to solve a problem logically. Arranging information together in a different way by merging the elements in a new pattern or an alternative solution. Here are an example of the test items found:

Write a short 'essay' about tragedy and comedy.

According to the finding presented, it is seen that Taxonomy Bloom (especially test of evaluating and test of creating) provides framework in developing test questions. It covers the high-order of thinking skill by requiring students to do (collecting fact, applying knowledge, making prediction, solving a problem or evaluating theory). So, taxonomy Bloom influenced the way a lecture designs the final assessment, and describes the result of learning. It is stated on the verb that implies the kind of exercise done by student and the kind of grading. On the test of creating, lecturers can give a straight monitoring on exercise and students' behavior through the completion of the test. The verbs used in the test of evaluating and test of creating describe the skill of complex thinking, including knowledge application on practical problem, analysis of competitive interpretation, and the creating of new knowledge or the alternative interpretation of a problem.

This study focuses on the data test items in the midterm test and final test exercises by using taxonomy Bloom theory. Thus this study describe the implementation of test of evaluating and test of creating which have been applied in midterm and final test by the EFL lecturers. This study proves that critical thinking as one vital aspect in higher education field is a must have competency for university graduates since they are demanded to be able to think critically in the world of work, this study supported by Kamlasi and Sahan (2018) who found that in the arrangement of teacher's test should distribute the domains of Bloom taxonomy.

Conclusion

Lecturers use test of evaluating by inquiring students to review based on criteria and standards. Test of creating requires students to be able to create ideas and new information using the material that has been studied before. Bloom Taxonomy indeed is able to describe the cognitive demand needed by lecturers in designing the learning and .the assessment. With 'the verbs' divided in each cognitive level, lecturers can give the meaning in the context made by other verbs which mean taking one purpose, but it might obtain different meaning of the created context.   It is suggested that the EFL lecturers who had employed Bloom's taxonomy verbs should select that verbs for the learning outcomes which cover all the details of test takers' level. The EFL lecturers requisite consider the contextual contents of learning of outcome.

References

Aghaei, K., Education, P., Rad, E. M., & Science, G. (2018). *On the Interconnection between Bloom's Critical Thinking Taxonomy & Listening Comprehension Performance of Iranian EFL Learners* . 22–31

Akib, E., & Arief M (2019). Assessment of Teaching in 21st Century. *Journal of Physics: Conference Series,   1 (1179).*

Anderson, Lorin W, and Krathwohl, D. R. (2001). *TAXONOMY for learning, teaching and assessing* (A Bridged; L. Anderson, D. Krathwohl, P. Airasian, K. Cruikshank, R. Mayer, P. Pintrich, ... M. Wittrock, Eds.). David McKay Company Inc.

Bloom, B. (1956). *Taxonomy of Educational Goals*. Canada: Longman LTD.

Cohen, Louis & Lawrence Manion (2007). Research Methods in Education. New York: Routledge

Ebadi, S. (2016). Exploring Bloom's Revised Taxonomy of Educational Objectives in TPSOL Textbooks. *Journal of Teaching Persian to Speakers of Other Languages*, *5*(1), 65–93.

Eber, P. A., & Parker, T. S. (1999). *Assessing Student Learning : Applying Bloom ' s Taxonomy Bloom ' s Taxonomy*. (1956), 45–54.

Ekahitanond, V. (2013). *Promoting university students ' critical thinking skills through peer feedback activity in an online discussion forum*. *59*(2), 247–265.

Kamlasi, Imanuel & Anselmus Sahan. (2018). Journal of English Language Literature and Teaching. Descriptive Analysis on English Test Items based on the Application of Revised Bloom's Taxonomy, Vol 2, p 203-210

Köksal, D., & Ulum, Ö. G. (2018). JOURNAL OF LANGUAGE AND LINGUISTIC STUDIES Language a ssessment through Bloom ' s Taxonomy. *Journal of Language and Linguistic Studies*, *14*(2), 76–88

Krathwohl, D. R. (2014). *Revision An of Overview Bloom ' s Taxonomy :41*(4), 212–218.

Ledward, B. C., & Hirata, D. (2011). An Overview of 21st Century Skills. Summary of 21st Century Skills for Students and Teachers, by Pacific Policy Research Center, Honolulu: Kamehameha Schools-Research & Evaluation.

Leslie, C. (2016). *Anderson and Krathwohl Bloom ' s Taxonomy Revised*.

Lindemann, D. J., & Tepper, S. J. (2012). Painting with Broader Strokes: Reassessing the Value of an Arts Degree--Based on the Results of the 2010 Strategic National Arts Alumni Project. Special Report 1. *Strategic National Arts Alumni Project*

Mbato, C. L. (2019). Indonesian EFL Learners ' Critical Thinking in Reading : Bridging the Gap between Declarative , Procedural and Conditional Knowledge. *HUmaniora*, *31*(1), 92–101.

Paul, R., & Elder, L. (2008). *The Miniature Guide to Critical Thinking*.

Setyowati., & Susanto (2019). *Promoting University Students' Critical Thinking Skills Through Implementation of the Revised Bloom's Taxonomy in Language Assessment: A Theoretical framework Perspective*.181-184 (SoSHEC 2019)

Stanny, Claudia J. (2016), Reevaluating Bloom's Taxonomy: What Measurable Verbs Can and Cannot Say about Student Learning. 1-12 (MDPI Education Series)

Thomas, T. (2017). *Developing First Year Students ' Critical Thinking Skills*. *7*(4). https://doi.org/10.5539/ass.v7n4p26

Wilson, K. (2016). Critical reading , critical thinking : Delicate scaffolding in English for Academic Purposes (EAP). *Thinking Skills and Creativity*, *22*, 256–265. https://doi.org/10.1016/j.tsc.2016.10.002

Zuhri, F., & Susanto. (2018). *A Study on the Competence of English Subject at Junior Middle School Having Low and High Order Thinking Skills*. *108* (SoSHEC 2017), 60–64.