

**THE AUTHENTICITY OF ENGLISH LANGUAGE ASSESSMENT FOR
THE TWELFTH GRADERS OF SMK(VOCATIONAL HIGH SCHOOL)
NEGERI 4 SURAKARTA**

Nurul Hidayati

*An English Teacher of SMKN 4(State Vocational High School) Surakarta Indonesia
nurulhidayatik4@gmail.com*

Abstract :

The new curriculum applied in Indonesia recommended the use of authentic assessment. This study deals with the authenticity of English assessment. This research aims at finding empirical facts about what type of assessments are used, describing the authenticity of the assessments, showing teachers' understanding to the assessment system and identifying problems encountered in applying authentic assessment. The method used is qualitative research, and the data was collected by observation, interview and content analysis. The findings are firstly, the teacher used three types of assessment : (a) Formative Test covering affective-cognitive- psychomotoric aspects (b) Mid Term Test (c) Semester Test. Secondly, the Formative Test had high level of authenticity, while the Mid Term Test and Semester Test had low authenticity. Thirdly, the scoring system used was the Criterion Reference Test which refers to Minimum Mastery Criteria. (4) There were at least four problems in applying authentic assessment : (a) the teacher felt overburdened with too many assessment formats (b) the inconsistency in regulation cause confusion for teachers (c) insufficient learning facilities caused ineffectivity (d) insufficient IT system needed improvement. The school had been doing efforts to overcome the problems.

Keyword: *authenticity, assessment, and authentic assessment*

INTRODUCTION

Curriculum and teaching methods continue to change (Bagley, 1995:220). The educational regulation in Indonesia is about the change of the National Educational Standard states that the curriculum needs to be reconstructed to improve the quality and competitiveness of the nation. It is also aimed to keep in pace with the development of society either local or global.

The need for alternatives to traditional assessment instruments is growing. Alternative assessment consists of any method of finding out what a student knows or can do that is intended to show growth and inform instruction and is an alternative to traditional forms of testing, namely, multiple choice test. (Stiggins 1991 in O'Malley and Pierce, 1996,p.1). This alternative assessments are typically authentic assessment because it represents classroom and real life settings.

This study focuses on the authentic assessment which is used in assessing integrated English language skill as it is implemented in vocational high school. The implementation of authentic assessment in vocational school is based on the Guidelines of Curriculum 2013 Implementation issued by the Ministry of Education and Culture in Indonesia.

Based on the background of this study, the problems are (1) What types of assessment are used by the teachers in English Language Assessment? (2)How is the authenticity of the assessments used by the teacher ?(3) How do teachers use the assessment to score the students' learning outcome? (4) What problems are encountered by the teachers in applying authentic assessment for English language ?

The objectives of this study are as (1) to describe what types of assessment are used by the teachers in English language assessment,(2) to know how the authenticity of the assessments used by the teacher is,(3) to know how teachers use the assessment to score the students' learning outcome,(3) to identify the problems encountered during the application of the authentic assessment and the solution.

The finding of this study can be used as profound study to explore more about the implementation authentic assessment to assess different language skill. The findings of this research are also expected to be useful for teachers at providing reference about the types of authentic assessment and its implementation based on curriculum 2013 and reference to conduct further research on authentic assessment in English Language Assessment.

The implementation of new curriculum in Indonesia absolutely requires practices of authentic assessment. This includes the assessment on the aspects of cognitive , performance and attitude. As the ultimate objective of the English learning is for International communication, speaking skill becomes priority.

There are several previous studies which are relevant to the practice of authentic assessment. Most of pedagogical researches on authentic assessment were specialized in certain kind of form, e.g, portfolio or peer assessment (Bagley , 1995; Chen 2006 ; White, 2009, Paragae et al 2103, Alabdewahab (2002)). Other researches viewed authentic assessment from the perspective of the teachers (Azis, 2014 ; Nurgiyantoro, 2011). The other one view the authentic assessment from he perspective of teachers and students. Of all the researches only one study concerned with speaking skill, but it specialized in public speaking (White, 2009). The kind of education under study varied from formal and the level also varied from informal education, junior, senior, university.

The position of the current study is mostly inspired by the study conducted by Nurgiyantoro (2011) reporting a surprising fact that in general, the teachers do not understand and do not carry authentic assessment in language learning. However, it was the facts in 2011 when authentic assessment was not as popular as it is today in Indonesia due to the implementation of Curriculum 2013. The more recent research conducted by Komarudin (2014) showed that authentic assessment was not optimally implemented because some teachers still had difficulties in implementing authentic assessment. The writer is interested to conduct a similar research in the scope of English language teaching in Vocational High School of Tourism, SMK Negeri 4 Surakarta the twelveth Graders, to analyze whether or not the assessment used to assess English as Foreign Language reflects authentic assessment. This idea is relevant to the fact that authenticity of assessment, as Paragae et al. (2013,p.1) found in his research, gives high contribution to the students' achievement.

Gronlund (2003:14) defines assessment as “the various methods used to determine the extent to which students are achieving the intended learning outcomes of instruction”. That is, unlike tests, the results of assessment do not need to be reported all the time, and furthermore, they are not necessarily reflected in the final grades. Thus, the term assessment is used to incorporate a wide range of methods for evaluating student performance and attainment, including formal testing. Specific terms (i.e., standardized tests, formal tests, or school- administered tests) are used in cases where the discussion focuses on specific forms of assessment (Lee, 2010,p.8).

Based on the function, assessment can be identified into formative and summative assessment. Formative assessment evaluates students in the process of forming · their competencies and skills with the goal of helping them to continue that growth process. While summative assessment aims to measure, or summarize, what a student has grasped, typically occurs at the end of a course or unit of instruction. Final exams in a course and general proficiency exams are examples of summative assessment (Brown, 2004,p. 6).

Armstrong, 1994, and Bailey, 1998 : 207 in Brown 2004, p.13) make distinction between traditional and alternative assessment. Traditional assessment has the characteristic of : one-shot, standardized exams, timed, multiple choice format, de-contextualized test items, Scores suffice for feedback, norm-referenced scores, focus on the "right" answer , summative, oriented to product, non-interactive performance , fosters extrinsic motivation. While alternative assessment has the characteristic of : continuous long-term assessment, untimed, free-response format, contextualized communicative tasks, individualized feedback and wash-back, criterion-referenced scores , open-ended creative answers, formative , oriented to process, interactive performance, fosters intrinsic motivation

The earliest reference to authentic tests is likely that made by Archibald and Newman in 1988, in a book critical of standardized testing, that sought to

promote assessment centered on meaningful real-world problems or tasks. Assessment is authentic when it measures products or performances that “have meaning or value beyond success in school” (Newman, Brandt & Wiggins, 1998 in Frey, et al , 2012,p. 2).

Authentic assessment is a term to describe a variety of alternative assessment methods that allow the student to demonstrate his ability in completing tasks and solving problems. Simultaneously, express knowledge and skills in a way to simulate situations that can be encountered in the real world outside the school environment (Hymes, 1991).

O'Malley and Pierce (1996,p.11) mention the types of Authentic Assessment are: Oral Interview, Story or Test Retelling, Writing Samples, Projects/Exhibitions, Experiments/Demonstrations, Constructed-Response Items, Teacher Observations, and Portfolios.

The followings are clarification about some types of authentic assessment :

1) Self Assessment and Peer assessment

Involving students in their own assessment is critical. By reflecting on and assessing their own work and that of their peers, students get their opportunity to apply criteria to work samples and to set the learning goals (Brown, 1996:66) . Peer assessment in language learning is most frequently associated with peer reviews of the writing process (Anderson, 2012 ,p.189).

2) Performance Assessment

Performance tasks are hands-on activities that require students to demonstrate their ability to perform certain actions. This category of assessment covers an extremely wide range of behaviors, including designing products or experiments, gathering information, tabulating and analyzing data, interpreting results, and preparing reports or presentations (M.Stecher et al, 2000 ,p. 27)

Performance assessment requires students to accomplish complex or significant tasks, while bringing to bear prior knowledge, recent learning and relevant skill to solve realistic or authentic problem (O' Malley 1996 ,p. 4-5)

3) Portfolio Assessment

According to Arter & Spandel (1992), portfolio is a purposeful collection of student work that exhibits to the student, or others, her efforts or achievement in one or more areas. Paulson, Paulson and Mayer (1991, p. 60) define that portfolio is a purposeful collection of student's work that exhibits the student's efforts, progress and achievement in one or more areas. The collection must include student participation in selecting contents, the criteria for selection, the criteria for judging merit and evidence of student self-reflection.

4) Observation : Anecdotal Records and Developmental Checklists.

Anecdotal Records and Development Checklists are informal observation about what students are learning, how students are responding to instructions, or any other students behaviors' action or reaction that might provide teachers with some insight (Rhodes and Nathenson-Mejia 1992 in Winograd and Perkins (1996 : 1-8:5).

5) Student Projects

Senior projects are distinct from written assessments and performance tasks because they are cumulative, i.e., they reflect work done over an extended period rather than in response to a particular prompt. It has three components: a research paper, a product or activity, and an oral presentation, all associated with a single career-related theme or topic (Stecher, B. M., & Bohrnstedt, G. 2000 ,p. 29).

6) Interviews

Interviews can assess student knowledge of many topics. Asking a student why he/she selected a certain play while coaching a team can reveal understanding of offensive or defensive strategies

According to Brown (2004, p.19-29) , there are five cardinal criteria for testing a test : practicality, reliability, validity, authenticity and washback. The Authenticity of Assessment in English Language Teaching was discussed by many expert among them are Brown and Hudson. They stated the characteristics of alternative assessments can be summarized as follows : (1) require students to perform, create, produce or do something, (2) use real-world context or simulations (3) are nonintrusive in that they extend day to do classroom activities (4) allow students to be assessed on what they normally do in class everyday (5) use task that represent meaningful instructional activities (6) focus on process as well as products (7) Tap into higher level thinking and problem solving skills (8) provide some information about both strength and weaknesses of students (9) are multiculturally sensitive when properly administered, (10) Ensure that people not using machines do the scoring human judgment (11) Encourage open disclosure of standard and rating criteria, and (12) Call upon teachers to perform new instructional and assessment role.

According to the Standard of Assessment issued by the Ministry of Education and Culture of Republic Indonesia (2014) , authentic assessment is a form of assessment that requires students to show attitude, use the knowledge and skills gained from learning in performing tasks in real situations. That is to show that there are four characteristics of authenticity in an assessment , they are : 1) requires students to show attitude; 2) requires students to use knowledge; 3) requires students to use skill 4) requires students perform tasks in real situation.

When a teacher designs a test, one must consider how the test will be scored and graded. The scoring plan reflects the weight of each section or item in each section (Brown 2004 ,p.61). Scoring rubrics are rating scales as opposed to checklists-that are used with performance assessments. They are formally defined as scoring guides, consisting of specific pre-established performance criteria, used

in evaluating student work on performance assessments. Rubrics are typically the specific form of scoring instrument used when evaluating student performances or products resulting from a performance task (Mertler, 2001,p.2)

One important thing after scoring is defined , student scores must be converted to “grades.” The process of converting rubric scores to grades or categories is more a process of logic than it is a mathematical one (Trice , 2000 in Mertler, 2001,p. 5).

Scoring and Grading System in Indonesian Education Based on Curriculum 2013. The scoring system uses the principles of authentic assessment. The scoring covers three aspects , i.e, affective aspect, cognitive aspect and.

The principles of authentic assessment from the experts can summed up into ten : (1)requires students to show attitude, knowledge and skill (2) requires students perform tasks in real situation and have value and meaning beyond the classroom (3) involves metacognition process(4) use higher-order thinking skills(5) focus on process as well as products (6) require students to perform, create, produce or do something, (7) measure the expectations, respect, and extent of inclusion of all students in the learning process(8) provides some information about both strength and weaknesses of students (9)uses task that represent meaningful instructional activities(10) does not use machines to do the scoring

RESEARCH METHOD

The type of research used in this study was qualitative research. The research was conducted for six months, starting from the first week of July 2015 to the forth week of December 2015 at SMK Negeri 4 Surakarta. The research object was the authenticity of English Language assessment for the twelfth grade students of SMK Negeri 4 Surakata in the academic year of 2015/2016. The subject of the research was two teachers and one school administrator . They are

purposively assigned as the informants, because as compared to other teachers, they have implemented new curriculum longer than the three others. They have implemented the new curriculum for more than two years.

The data source of this study are informant , documents and event. The techniques of collecting the data used in this study are (1) observation to get the data about the school profile, the learning circumstances in SMK Negeri 4 Surakarta, the profile of the teachers and students, and classroom language learning activities (2) interview which was based on the purposeful sampling to get information from the human resources, in this case, teachers and the school administrator, the vice principal of curriculum (3) content-Analysis concerned with the teachers' lesson plan that the teachers prepared before the teaching The lesson plan stated the assessment planning including technique, instrument and the scoring rubric. The researcher also analyze the students' work or products to get the data about the authenticity of the assessment.

To assess the accuracy of the findings as best described by the researcher and the participants it need data validation. In validating the data, internal validity was used in this research to ensure the degree of accuracy of the design of the study with the results . The writer used Miles and Huberman's technique of analyzing data. The procedures can be described as (1) Data reduction (2) Data displays (3) Conclusion drawing/verification.

RESEARCH FINDINGS AND DISCUSSION

The writer describes the findings about the authenticity of tests applied in the English language assessment based on the research questions formulated in chapter one. The findings are presented in sequence started from types of assessment are used, the authenticity of each type of the test, how do teachers score the students' learning outcome and what problems are encountered by the teachers.

1. The Types of Assessment Used by The Teachers in English Language Assessment

Based on the interview with two teachers and one school administrator, supported by the data from the document study, it was found that the assessment used by the teachers in English language assessment during the odd semester included three types. The three types were : a) Formative Test including three aspects, namely, affective, cognitive and psychomotoric. The types of assessment used in each aspect are : observation for affective aspect; written test for cognitive aspect; performance test, product and portfolio for psychomotoric aspect; b) Mid Term Test; and c) Semester Test.

Formative test is an assessment which is done by the teachers during and after completing one learning content. In the formative test, the test instrument and administration were prepared by the teachers and the scoring was also done by the teachers. The assessment covered the aspect of affective, cognitive and psychomotoric. Each can be described as follows :

The types of assessment used in assessing affective aspect is observation. Observation is a part of assessment which provides data about students behaviour while they interact with their surrounding during the learning process. The observation was done by teachers while the teaching learning was going on in the classroom. The teacher used checklist to observe the students' behaviour. The items being assessed was matched with the basic competencies stated in the syllabus.

The types of assessment used in assessing cognitive aspect was written test. The written test used in the formative test was a paper test containing five to ten items and used various formats such as short answer, completion, fill in the gap, paraphrasing and jumbled sentences. The students sheet was then corrected and scored by the teacher of each class. Here is an example of formative test taken from the teacher's document.

The types of assessment used in assessing psychomotoric aspect were Performance test, Product and Portfolio.

Performance assessment is an assessment in which students are exposed with complex and significant tasks which require students to demonstrate their ability or to perform certain actions. Criteria of expected behaviour and the scoring system are well-defined to keep the objectivity of the assessment.

Performance test was used by the teachers to perform presentation of theme dialog/conversation, retelling, expressing opinion and presenting the result of discussion. Performance test could be done either individually or in group. Here is a picture which was taken during the class observation.

The scoring of product includes three stages, namely: preparation stage ; stage of production (process); stage of product assessment (appraisal) . The teachers used this type of test to assess students' writing product. The writing product found in the teachers' document was Caption writing. The captions were made in groups of three to four students.

Portfolio is a purposeful collection of student work , it is a record of the students' process of learning and what the students have learned . It is an important document because it provides data about students' skills in certain area at a certain time. Portfolio assessment was done during the month of September to October 2015. One of the students' portfolio found in the teachers' document was the students' collection of job vacancy in English language, along with the application letter and resume.

The Mid Term Test is an assessment conducted to assess learning contents taught within the time frame of the first half of a semester. It measures students' competencies against learning content taught within that time frame. The instruments and administration of mid term test were prepared by the school teachers who were assigned by the headmaster to be the Committee of Mid Term Test. The assigned teachers are in charge of preparing assessment instrument for

mid term test for one grade. The person who was in charge for grade XII will be different from those for grade XI or grade X.

The types of assessment used Mid Term Test was written test. The written test conducted belonged to the traditional type of test which had the characteristic of pencil and paper test, timed, de-contextualized test items, non-interactive performance .

However; the scoring was done by the class teacher. Here is an example of formative test taken from the teacher's document. The mid term test was a written test containing fifteen to twenty items and used various formats such as short answer, completion, fill in the gap, paraphrasing and jumbled sentences. The students sheet was then corrected and scored by the teacher .

The semester test is a written assessment conducted to assess learning content taught within the time frame of one semester. It measures students' comprehension against key competencies at the end of one semester. The key competencies cover four skills including speaking, reading, writing and listening, which are conducted at the local government level. The test covers all basic competencies taught within the time frame of six months and conducted in the form of written test and follow the pattern of national examination.

English Teachers Association in the local government of Surakarta conducts the summative test for all SMKs in Surakarta with the same test, schedule, formats and time allocation.

The test instruments for three subjects being tested in National Examination, Math-English-Bahasa Indonesia, were administered by the the board of vocational school principals in Surakarta. The test instruments were prepared by a selected team consisting teachers from different school. The scoring used computerized system, under the management of the school principal association.

The semester test belonged to the traditional type of test which had the characteristic of standardized exams, pencil-paper test, timed, multiple choice format, de-contextualized test items, focus on the "right" answer, non-interactive performance .

The students need to blacken the options provided (A, B, C, D). The sheet requires students to present their personal data such as name, signature, identification, test number, school name, the subject and the date of test. All information on the sheet must be written using 2B pencils because it would be processed using computer. The result of the test then was used as one of components in deciding the final score for student. The result was also used by the local government for mapping the achievement of schools under the supervision of the local government in terms of the target of achievement in National Examination.

2. *The authenticity of the assessments used by the teacher*

The formative test applied by the teachers the following principles of authentic assessment such as : 1)requires students to show attitude, knowledge and skill (2) requires students perform tasks in real situation and have value and meaning beyond the classroom (3) use higher-order thinking skills(5) focus on process as well as products (6) require students to perform, create, produce or do something, (7) provides some information about both strength and weaknesses of students (8)uses task that represent meaningful instructional activities(9) does not use machines to do the scoring. Out of ten principles, there were eight which matched the criteria. It can be inferred that the formative test had high level of authenticity.

The Mid Term Test belonged to non-authentic assessment because it had very low level of authenticity. Out of ten principles, there were only two matched the criteria. It was similar to the Semester test which matched only one characteristic of authentic assessment.

Regarding the authenticity of the assessment used by the teacher and the ten principles of assessment authenticity, it can be inferred that the teachers used both authentic and non-authentic test in assessing the students learning outcome. The authentic one refers to formative and non-authentic ones refer to Mid Term Test and Semester Test.

3. The Scoring System Used by Teachers to Score the Students' Learning Outcome.

In determining the final score, the teachers referred to the regulation stated in The Decree of The Minister of Education and Culture of Republic Indonesia Number 104 of the Year 2014 about the Assessment by Teachers in Elementary and Secondary Education.

The scoring system referred to the Criterion Reference Test. It means that the student's score obtained from an assessment result, both formative and summative, was not compared to scores of other students, but compared to the minimum mastery criteria of the competencies required. For those who had not reached the criterion, they were given the opportunity to participate in remedial learning done after an assessment. The Minimum Mastery Criteria for English in the academic year of 2015/2016 is 2,67 for Cognitive Score and Psychomotoric Score. While the Minimum Mastery Criteria for affective score is B (Good).

In determining the final score for the school report at the end of semester, the teachers used the following rules: (a) the final score for the affective is obtained from the value of modes or the most frequently value appears (b) the final score of cognitive was drawn from the value of mean or the average of all scores (c) the final score for the psychomotor is taken from the optimum value or the highest value reached.

4. The Problems Encountered by the Teachers in Applying Authentic Assessment

The problems encountered by the teachers in applying authentic assessment can be described as follows.

- a. Teachers felt overburdened with too many assessment formats

The teacher had to prepare many teaching formats for the assessment. Each type of test had its own format. During the odd semester the English teacher of grade XII had to finish 10 learning content with the time allotment of 40 hours. The teachers had to make assessment of three aspects, affective-cognitive-psychomotoric, for each learning content. It means that in one semester she had to make 30 assessments. If she handled 12 classes she would do 36 assessments. That was very overburdening teachers.

b. The inconsistency in educational regulation made teachers confused

Curriculum 2013 was applied since July 2013, but the changes in the educational regulation dealing with its implementation still continued until 2016. The regulation dealing with assessment had changed at least three times. The first one is stated in the Decree of Minister of Education and Culture No 66 the Year of 2013 about the standard of assessment, then the Decree of Minister of Education and Culture No 81A the Year of 2014, and again it was revised by the Decree of Minister of Education and Culture No 104 the Year of 2014. The changes affected the scoring system, and teachers as a practitioner were mostly affected by the changes.

c. Insufficient learning facilities caused inefficiency

The implementation of authentic assessment need more learning facilities to support. The availability of LCD projector in the classroom is vital. All the classroom were facilitated with LCD, nevertheless some them are not in good condition. The availability of internet access for students could not meet the need of all students.

d. Insufficient IT system needed improvement

The scoring system used in Curriculum 2013 is complicated. It was almost impossible for teachers to do all the assessment administration manually without the help of computerized system. Although most of teachers are computer literate,

they still have problem with the scoring system. The following is the part of interview with the Vice Principal of Curriculum

To overcome the problems encountered by the teacher, some efforts were made by teachers and school administrators .

Compared with the findings in the study by Nurgiyantoro and Suyata (2001) is that in the previous study it was found that in general, teachers do not understand and did not make authentic assessment in language learning. While the findings in the current study revealed that most teachers have good understanding in authentic assessment. It can be seen from the variety of assessment which they apply in assessing the students' learning outcome.

This is in line with the standard of assessment issued by the Minister of Education and Culture No 104 the Year of 2014 about Assessment of Students' Outcome by Teachers in the Secondary Education. The school applying curriculum 2013 must use the assessment covering the aspect of affective, cognitive and psychomotoric.

This is also in line with the statement by Hymes (1991) that authentic assessment is a term to describe a variety of alternative assessment methods that allow the student to demonstrate his ability in completing tasks and solving problems. Simultaneously, express knowledge and skills in a way to simulate situations that can be encountered in the real world outside the school environment .

The current study has similar finding compared to the previous study conducted by Paragae et al. (2013) which found the the assessment made by teachers had high authenticity level.

This is in line with the principles of authentic assessment stated by some experts such as Hymes (1991) , Newmann and Wehlage (2007), McAlister (2000), Brown and Hudson (2004), and The Standard of Assessment issued by the Ministry of Education and Culture of Republic Indonesia (2014).

The scoring system used by the teachers are in line with the regulation about scoring and grading system applied in Indonesia based on Curriculum 2013 which is explained in detail in the Decree of The Indonesian Minister of Education and Culture Number 104 / 2014.

Compared with the finding in the study conducted by Komarudin (2015) there are some similarity and differences. The similarities are in the sense of types of assessment used by the teachers. The difference lies in the problems encountered by the teacher.

CONCLUSION

The teacher used three types of assessment, namely, Formative Test which covers affective-cognitive-psychomotoric aspects, Mid Term Test and Semester test. Observation used in assessing affective aspect, written test was used in assessing cognitive aspect and performance-project-portfolio in assessing psychomotoric aspect. Written test was also used in Mid Term and Semester Test.

The Formative Test used by the teachers has high level of authenticity, while the Mid Term Test and Semester Test had very low authenticity. In other words, teachers used both authentic and non authentic assessment in assessing students' outcome.

The scoring system used the Criterion Reference Test which refers to Minimum Mastery Criteria. The Minimum Mastery Criteria for affective aspect was B, and for cognitive and psychomotoric aspect was 2,67.

There were at least four problems in applying authentic assessment, namely, (1) too many assessment formats (2) the inconsistency in educational regulation (3) insufficient learning facilities caused inefficiency (4) insufficient IT system need improvement. The school had been doing efforts to overcome the problems.

By recognizing the authenticity in the assessment, some pedagogical implication can be provided based on the findings of this study :

Authentic assessment is considered to be an alternative assessment which can measure the students competency in many aspects and it is proved that it gives the students meaningful learning experience, so the school elements should support its implementation.

In developing assessment teachers should always refer to the principles of authentic assessment and the educational regulation, response the changes positively and continually improve themselves to be more professional in assessing students

The teacher should make good preparation in before conducting assessment so that the scoring can be done easily and effectively.

It is important for the school management to support the implementation authentic assessment by providing good facilities , facilitating teachers to improve competency in IT and assessment, and providing good assessment management system.

REFERENCES

- Alabdelwahab, Sharif Q., (2002), *Portofolio Assessment : A qualitative investigation of portofolio self-assessment practices in an intermediate EFL Classroom, Saudi Arabia*, Ohio State University
- Anderson, Neil J, (2012). *Student Involvement in Assessment : Healthy Self-Assessment and Effective Peer Assessment*. The Cambridge Guide to Language Assessment. Edited by Christine Coombe, Peter Davidson, Barry O’Sullivan, Stephen Stoyloff. Cambridge University Press. New York.
- Arter, J. A., & Spandel, V. (1992). *Using portfolios of student work in instruction and assessment*. Educational measurement: Issues and practice, 11(1), 36-44.
- Azis, A. (2014). *Indonesian junior secondary school teachers' conceptions of assessment: A mixed methods study*.

- Bagley, Pamela Bryan,(1995). "*Authentic Assessment: How Do Portfolios Fit the Picture?*" . UNF Theses and Dissertations. Paper 220. Retrieved April, 2015 from <http://digitalcommons.unf.edu/etd/220>)
- Brown, H. D. (2004). *Language assessment: Principles and classroom practices*. Allyn & Bacon.
- Chen , Yuh-Mei. (2006), *EFL Instruction and Assessment with Portfolios: A Case Study in Taiwan*, *The Asian EFL Journal Quarterly*, March 2006 Volume 8 Issue 1.
- Gronlund, N.E. (2003). *Assessment of Student Achievement*. 7th ed. Custom edition for the University of Alberta.
- Hymes, D. L. (1991). *The Changing Face of Testing and Assessment: Problems and Solutions*. AASA Critical Issues Report. American Association of School Administrators, 1801 North Moore St., Arlington, VA 22209-9988.
- Komarudin. (2015). *Penilaian Autentik pada mata Pelajaran Pendidikan Agama Islam dan Budi Pekerti (Studi Analisis Kurikulum 2013 Kelas VIII Semester 1 di SMP Negeri 4 Yogyakarta Tahun Pelajaran 2014/2015)*. Unpublished thesis. Post Graduate Program of UIN Sunana Kalijaga
- Lee , Sujin.(2010). *Current practice of Classroom Speaking Assessment in Secondary Schools in South Korea*, A Thesis Submitted in Partial Fulfillment of the Requirement for the Degree of Masters of Arts (MA) in Applied Linguistics, The University of Queensland.
- Mertler, Craig A. (2001). *Designing scoring rubrics for your classroom. Practical Assessment, Research & Evaluation*, 7(25). Retrieved April 22, 2015 from <http://PAREonline.net/getvn.asp?v=7&n=25> .
- M. Stecher, Mikala L.Rahn, Alen Ruby, Martha Naom Il, Abby Robyn, Brian Ward. (2000). *Using Alternative assessment in Vocational Education*. National Center for Reseach in Vocational Education University of California . Berkeley.
- Newman, Fred M., King, M.Bruce, Carmichael, Dana L. (2007). *Authentic Instruction and Assessment: Common Standards Rigor and Relevance in Teaching Academic Subject*, *The Iowa Department of Education*.Des Moines.

- Nurdiyantoro, B. (2011). *PENGEMBANGAN MODEL ASESMEN OTENTIK DALAM PEMBELAJARAN BAHASA*. Cakrawala Pendidikan Edisi November 2009.
- O'Malley, J. M., & Pierce, L. V. (1996). *Authentic assessment for English language learners: Practical approaches for teachers*. Reading, Mass: Addison-Wesley Pub. Co
- Paragae, I G.A.P.N.S.; Marhaeni, A.A.I.N.; Paragae, N.Dantes. (2013). , “Teacher Made Assessment Authenticity in Senior High Schools and Its Contribution to Students’ English Achievement”, e-Journal Program Pascasarjana Universitas Pendidikan Ganesha Program Studi Pendidikan Bahasa Inggris Volume 1
- Paulson, F. L., Paulson, P. R., & Meyer, C. A. (1991). What makes a portfolio a portfolio. *Educational leadership*, 48(5).
- Stecher, B. M., & Bohrnstedt, G. (2000). *Class size reduction in California: The 1998-99 evaluation findings*.
- White, Eddy , (2009). *Student Perspectives of Peer Assessment for Learning in a Public Speaking Course* , Asian EFL Journal – Professional Teaching Articles. Vol. 33. January
- Winograd, P. and F. D. Perkins. (1995). *Authentic assessment in the classroom: principles and practices*. In *A handbook for student performance assessment in an era of restructuring*, edited by R. E. Blum and J. A. Arter. Alexandria, VA: Association for Supervision and Curriculum Development
- , (2014). *Lampiran Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia Nomor 104 Tahun 2014 Tentang Penilaian Hasil Belajar oleh Pendidik pada Pendidikan dasar dan Menengah*. Kementerian Pendidikan dan Kebudayaan Republik Indonesia. Jakarta