

CHAPTER II

REVIEW OF RELATED LITERATURE

This chapter presents and discusses about the concept of the assessment, the concept of authentic assessment, assessing speaking, authentic speaking assessment, problem of authentic speaking assessment, rubric of authentic speaking assessment.

2.1 The Concept of Assessment

Assessment is defined as gathering information about students then the gathered information is used to assistance teachers in the judgment process (Anderson, 2003:4). While Brown (2003:4) defines assessment as an ongoing process that encompasses a much wider domain. Whenever a student conduct performance such as responds to a question, or tries out a new word or structure, offers a comment, so on, the teacher subconsciously conducts an assessment on it.

When conducting classroom assessment, teacher has to consider principles of effective classroom. According to Marzano (2006:3), there are four principles of effective classroom assessment. They include:

1. Giving a clear picture of their development on learning goals and how they might improve to students are important in the feedback from classroom assessments.

2. Promoting students to improve should be done in the feedback on classroom assessments.
3. Formative in nature should be done in the classroom assessment.
4. Formative classroom assessments should be conducted regularly.

Next, teacher has to include several domains in conducting a classroom assessment. As Anderson (2003:26-43) states that assessment must include several domains. Those are classroom behavior, student's effort and student's achievement. Teacher collected information about negative behavior through monitoring classrooms' behaviors silently. Next, student's effort including the amount of student's spare time to spend in learning, student's level quality of effort showing in learning and students' reason on inadequate effort. Last student's achievement, in simple, reaching student's achievement can be gotten through assessing students' knowledge of everything that the teachers say in class or everything that the students read in their textbooks.

Then teachers, when planning classroom assessment, should decide the goal of the assessment because it is important for teacher to make many decision throughout the school day. As Russel & Airasian (2012:7-8) state that the goal of classroom assessment is to establish a classroom that support learning, to plan and conduct instruction, to diagnose student problems and disabilities, to place students, to provide feedback and incentives, and to judge and grade academic learning and progress.

Then, based on its process, assessment is distinguished into informal and formal assessment. Informal assessment is an assessment that occurs incidentally during instruction, it can use number of forms, teacher' unplanned comments and responses, the goal is coaching and giving feedback to the student about how to improve work. While formal assessment refers to systematic, planned and constructed exercises or procedures specifically designed to tap into a storehouse of skills and knowledge, giving teacher and student an appraisal of student achievement is the goal (Brown, 2003:5-6).

Next, based on its procedures, assessment is divided into two. They are formative and summative assessment. Formative assessment, in which the goal is coaching students to continue the growth process, is defined as evaluating students in the process of building their competencies and skills. The key of this assessment is transferring by the teacher, and acquiring by the student, of appropriate feedback on performance. While summative assessment, in whose goal is to measure or summarize what a student has grasped, is evaluating students that occur at the end of a course or unit of instruction (Brown, 2003:5-6)

Then, based on its method, assessment is distinguished into two. They are authentic and traditional assessment. Traditional assessment refers to forced-choice measures, which is asked students to select an answer or recall information to complete it. The examples are multiple-

choice tests, fill-in-the-blanks, true-false, matching and the like that have been and remain so common in education. They may be standardized or teacher-created and administered locally or statewide, or internationally (Mueller, 2014:1). While Wiggins (1990:1) point out on directly examine student performance on worthy intellectual tasks are the scope of authentic.

From several point of view above, it can conclude that assessment is an ongoing process that encompasses a much wider domain for gathering, synthesizing, and interpreting information about student that can be used to aid teachers in the decision making process. Then, when planning assessment, the teacher should set domain and goal of assessment. Next, assessment based on its processes divides into informal and formal assessment, assessment based on its procedures divides into formative and summative assessment, and assessment based on its methods divides into traditional and authentic assessment. In the end, from two assessment procedures above, the last point authentic assessment is the major that will be discussed in this study.

2.2 The Concept of Authentic Assessment

Authentic assessment, also known as performance assessment or alternative assessment refers to the systematic recording of developmental observations over time by familiar and knowledgeable caregivers about the

naturally occurring competencies of pupil in daily routines (Bagnato and Yeoh in Bagnato, 2007:27). Then Bagnato (2007:27-28) defines authentic assessment as a set of plan for studying the natural behavior of pupils. In order to get the data, direct observation and recordings, interviews, rating scales, and observed samples of natural and thematic play and pupil' day-to-day living skills is conducted.

While Wiggins (1990:1) states that the assessment is authentic when the examiner directly measures student's performance on worthy intellectual tasks. In addition, Wiggins (1993: 229) states:

"...Engaging and worthy problems or questions of importance, in which students must use knowledge to fashion performances effectively and creatively. The tasks are either replicas of or analogous to the kinds of problems faced by adult citizens and consumers or professionals in the field".

Next, authenticity of an assessment as the similarity between the cognitive demands the thinking required of the assessment and the cognitive demands in the reflection or simulation of a real-life situation that could confront students in their internship or future professional life (Gulikers et al 2004:69).

There are few reasons of using authentic assessment. As O'malley's (1996:17) views:

The purpose of authentic assessment with English language

learner (ELL) students can include identification, placement, reclassification, and monitoring student progress. The three first purposes involves extremely important (“high-stakes”) decision that affect whether or not ELL student receive special language-based instruction, the type of instruction, and the duration over which the instruction continues.

Next, the function of authentic assessment in the assessment is not for replacing the position of traditional assessment but to complement the traditional assessment. As Mueller (2014:1) illustrated in his silly example:

...a teacher does not have to choose between authentic assessment and traditional assessment. It is likely that some mix of the two will best meet your needs. To use a silly example, if I had to choose a chauffeur from between someone who passed the *driving* portion of the driver's license test but failed the *written* portion or someone who failed the driving portion and passed the written portion, I would choose the driver who most directly demonstrated the ability to drive, that is, the one who passed the driving portion of the test. However, I would *prefer* a driver who passed both portions. I would feel more comfortable knowing that my chauffeur had a good knowledge base about driving in which might best be assessed in a traditional manner and was able to apply that knowledge in a real context in which could be demonstrated through an authentic assessment.

Testing and credentialing programs in education today often use both multiple-choice and performance assessment. Incorporating both performance tasks and multiple-choice items into an assessment may be due in part to assessment practitioners’ conceptualizing multiple-choice items and performance assessment as part of an item format continuum (Fortune et al in Johnson, 2009:14).

2.2.1 Types of Authentic Assessment

Feur and Fulton in O'malley (1996:12) state that there are lots of types of authentic assessment, in some cases performance-

based task or assessment, used in classroom today. They includes oral interviews, story or text retelling, writing examples, experiments or demonstrations, constructed-response item, portfolios, teacher observation, and conducting exhibitions or projects. Then, O'malley (1996:12) explains the types of authentic assessment through below figure:

Figure 2.2.1 Types of Authentic Assessment (O'malley, 1996:12)

Assessment	Description	Advantages
Oral Interviews	Teacher ask student question about personal background, activities, reading, and interest	<ul style="list-style-type: none"> • Informal and relaxed context • Conducted over successive days with each student • Record observation on an interviews guide
Story or Text Retelling	Students retell main idea or selected details of text experienced through listening or reading	<ul style="list-style-type: none"> • Students produce oral response • Can be score on content or language components • Score with rubric or rating scale • Can determine reading comprehension, reading strategies, and language development
Writing Samples	Students generate narrative, expository, persuasive, or reference paper	<ul style="list-style-type: none"> • Students produce written document • Can be score on content or language components • Score with rubric or rating scale • Can determine writing process
Projects/ Exhibitions	Students complete project in content area,	<ul style="list-style-type: none"> • Students make formal presentation, written report, or both • Can observe oral or written products and thinking skills

	working individually or pairs	<ul style="list-style-type: none"> • Score with rubric or rating scale
Experiments/ Demonstrations	Students complete experiment or demonstrate use of material	<ul style="list-style-type: none"> • Students make formal presentation, written report, or both • Can observe oral or written products and thinking skills • Score with rubric or rating scale
Constructed-response item	Students respond in writing open-ended questions	<ul style="list-style-type: none"> • Students produce written report • Usually scored on substantive information and thinking skills • Score with rubric or rating scale
Teacher Observation	Teacher observes student attention, response to instructional materials, or interactions with other students	<ul style="list-style-type: none"> • Setting in classroom environment • Takes little time • Record observations with anecdotal notes or rating scales
Portfolios	Focused collection of student work to show progress over time	<ul style="list-style-type: none"> • Integrates information from a number of sources • Gives overall picture of student performance and learning • Strong student involvement and commitment • Calls for student self-assessment

2.3 Assessing Speaking

According to Luoma (2004:29), assessing speaking is giving opportunities to the test-taker or test-takers to perform talk by the task or tasks that we give them then examiner or examiners judge their performances.

Nunan in Luoma (2004:40) states that there are two task in assessing speaking in classroom. They are pedagogic or language-focused tasks and real life task. While language-focused tasks are created specifically for certain types of language use, real life task simulate language use outside the classroom. Then Luoma (2004:40) gives examples of each kind of tasks. In language-focused tasks, learner A is given a simple graph with a blue triangle, a red square and a black circle arranged diagonally across a page and told to instruct learner B to draw the objects in the right configuration on an empty page. Then the real life task puts examinees in their professional role while the examiners act as customers, patients, guests, or other likely people who might do interaction with the examinees in the test language in occupational context.

Then, assessing speaking include both knowledge and competence domains. As stated by Bachman (1990:86) that framework of communicative language ability is knowledge and language competence. The language competence includes accuracy, fluency, organizational competence, grammatical competence and so on.

2.4 Authentic Speaking Assessment

Authentic speaking assessment, in other word is assessing speaking authentically, is giving opportunities to the test-taker to perform worthy intellectual talk (Wiggins 1990; Luoma, 2004).

Teacher can design authentic speaking assessment through oral interviews, project or exhibitions, story or text retelling, experiment or demonstration (O'Malley and Valdez, 1996:11-13). In oral interviews, teacher can ask student about personal background, activities, reading and interests. Then, in story or text retelling, student is asked to retell main idea with their own word. Next, in project or exhibitions, student is asked to make formal presentation then present it. Last, in experiment or demonstration, student is asked to demonstrate the use of materials and produce oral presentation (O'Malley and Valdez, 1996:12).

These authentic speaking assessments have several advantages. As O'Malley and Valdez (1996:12) states that these kind of assessment tend to be informal and relaxed context, conductive over successive days with each student, can be scored on content or language components, can observe oral skills, and produce oral presentation.

2.5 Problem of Authentic Speaking Assessment

There are several problems on authentic assessments, including authentic speaking assessment. First, authentic tasks promote "ill-structured" challenges, the consequence is this assessment opens to multiple interpretations rather than easily solved (Wiggins, 1990:1). Next, this assessment needs more time to conduct and administered and needs more higher budget than traditional assessment needed (Brown, 2003:13-

14; Wiggins, 1990:2). Then, *in developing this assessment, many times it is very difficult to design tasks with the proper level of difficulty (Dugin , N:8). Next, performance assessment in developmental process of review, revision and tryout takes months and is costly (Ruiz-Primo and Savelson, 1996:1052).*

2.6 Rubric of Authentic Speaking Assessment

Rubric means a scoring scale used to assess student performance along a task-specific set of criteria (Mueller, 2014:1). In speaking, rating scale expresses how well the examinees can speak the language being tested (Luoma, 2004:59) while authentic assessments typically measures based on criterion-referenced. That is, a student's aptitude on a task is determined by matching the student's performance against a set of criteria to determine the degree to which the student's performance meets the criteria for the task. To measure student's performance against a per-determined set of criteria, a rubric, or scoring scale, is typically created which contains the essential criteria for the task and appropriate levels of performance for each criterion (Mueller, 2014:1)

Next, a rubric is consisted of two components. They are criteria and levels of performance. There are at least two criteria and at least two levels of performance in each rubric. For each criterion, the assessor applying the rubric can determine to what degree the student has met the criterion, such as the level of performance (Mueller 2014:1).

Then, the reasons of including level of performance on the rubric are for communicating clearer expectations, for getting assessment that more consistent and objective, and for giving better feedback. As Muller (2014:1) states bellow:

“... it is very useful for the students and the teacher if the criteria are identified and communicated prior to completion of the task. Students know what is expected of them and teachers know what to look for in student performance. Similarly, students better understand what good (or bad) performance on a task looks like if levels of performance are identified. In addition to better communicating teacher expectations, levels of performance permit the teacher to more consistently and objectively distinguish between good and bad performance, or between superior, mediocre and poor performance, when evaluating student work. Furthermore, identifying specific levels of student performance allows the teacher to provide more detailed feedback to students. The teacher and the students can more clearly recognize areas that need improvement.

This figure bellow is the example of oral rubric:

Figure 2.6.1 The National Certificate descriptive scale (National Board of Education in Luoma, 2004:61)

Score	Criteria
6	Speaks fluently with few if any non-native features, such as a foreign accent. Is capable of expressing even subtle nuances of meaning with precision, and also makes varied and appropriate use of idiomatic expressions. Is able to describe even a complicated topic and to include sub-themes in the description, to develop different viewpoints and to bring the presentation to an appropriate conclusion.

Score	Criteria
5	Speaks fluently without frequent obvious need to search for an expression. Delivery characterized by naturalness, coherence and appropriate length. Is able to present a clear and detailed description of even a complex topic. Can use idiomatic expressions and everyday expressions, and is able to express nuances fairly well.
4	Copes fairly well even in less familiar speech situations. Makes a distinction between formal and informal registers, at least to some extent. Is able to present and justify an opinion comprehensibly. Is able to talk about and describe sights, sounds and experiences. Is obliged only rarely to use circumlocutions in everyday communication because of inadequate language proficiency.
3	Copes with the most familiar speech situations and is able to take the initiative in Every day language-use situations. Speech may be quite slow but there are few unnatural pauses. Is comprehensible despite transferring native or foreign language structures and vocabulary to the target language. Pronunciation may clearly deviate from target language standards.
2	Copes with routine speaking situations that require a simple exchange of information. Nevertheless, the speaker's language proficiency considerably restricts the range of matters that can be dealt with. Successful communication of a message presupposes that the interlocutor is willing to help the speaker in forming the message. Pronunciation may deviate clearly from the target language norm, thus requiring special effort from the interlocutor and impeding successful communication.

Score	Criteria
-------	----------

1	Is able to ask and reply to simple questions dealing with immediate everyday needs. Can make use of simple polite forms. Copes with the very simplest speaking tasks, but communication is slow and very fragmented. Often obliged to resort to nonverbal means in order to be understood.
---	--

From several points of views above, it can conclude that rubric is a scoring scale, for measuring student performance, has criteria, which determined to what degree the student has met the criterion, and level of performance, which are for communicating clearer expectations, for getting assessment that more consistent and objective, and for giving better feedback.

2.7 Previous Study

Dugin (2007), the study is about authentic assessment in college English instruction. This research focused on the theoretical and practical applicability of authentic assessment in English language instruction. Descriptive qualitative data analysis was used in this research. In this study, authentic assessment is implemented in two forms, one is process assessment, and the other is outcome-based assessment.

The differences of Duqin's research compared to my thesis are as follows: first, Duqin's thesis focused on English language instruction while my thesis focused speaking assessment. Second, Duqin's thesis

investigated wider domain of skills including reading, writing, reading and speaking skills while my thesis focused on speaking skill. Last, In the Duqin's study, authentic assessment is implemented through process assessment including formal and informal formative assessment and outcome-based assessment while in my study, the authentic assessment is implemented through formal formative assessment and summative assessment.