BEYOND TEST: ALTERNATIVES IN ASSESSMENT (A GLANCE OF THOUGHT FOR TECHNIQUES, CHALLENGES, AND OPPORTUNITIES DUE TO THE IMPLEMENTATION OF 2013 CURRICULUM)

Ni Luh Putu Eka Sulistia Dewi
Universitas Pendidikan Ganesha, Singaraja, Bali

Abstract: The changes of curriculum influences the assessment procedure implemented in elementary and high schools in Indonesia education system. Traditional test is no longer becomes the primary assessment instrument since the focus of the assessment is on the students' competence. Therefore, alternatives in assessment need to be put into consideration. The implementation of the alternatives assessment, such as portfolio, journal, and self and peer- assessment will bring their own challenges, and to face the challenges teachers need to broaden their insight on the opportunities that embedded on each alternative assessment techniques.

Keywords: Assessment, Alternative in Assessment, Portfolio, Journal, Self-Assessment, Peer-Assessment

Assessment is a popular and sometimes misunderstood term in current educational practice. Testing, assessing, and evaluating commonly considered as synonymous terms, but they are not. Brown (2004:4) justifies that tests prepared administrative procedures that occur at identifiable times in a curriculum. They are a set of assessment; they are certainly not the only form of assessment that a teacher can make. Assessment, on the other hand, is an ongoing process that encompasses a much wider domain. Johnson and Johnson (2002:2) emphasize that assessment continually collects information about the quality or quantity of a change in a student, group, class, teacher, school, or administrator, while evaluation is conducted occasionally; and the information from the assessment can be used to make evaluation. Tognolini and Stanley (2011) provide examples of assessment, such as classroom tests, checklists, practical work, project work, and when teachers ask a question in class they are assessing, and also when they observe what is going on in class they are also assessing. They are not as formal in their structure but they provide information that is just as pertinent and relevant about a student as the more highly structured means of collecting information is implemented.

In the current classroom instruction, traditional assessment, such as test, still becomes the most favorable language assessment technique used by the teachers no matter what competence being assessed, even though it has been viewed as somewhat unpopular area of language teaching and learning. Frequently, there is miss match between the competence and the assessment technique implemented. For instance, in assessing the students' competence in using language expression in speaking skill, the students are asked to write a dialogue in an essay test or choosing the correct expression from a multiple choice test. It does not provide evidence for the students' competence since there is no demonstrations of achievement involving actual performance (Johnson&Johnson, 2002:6). Moreover, students have sometimes viewed tests as unfair measure of their language ability and have feared that they would not perform well on them. Teachers have sometimes felt reluctant to construct classroom tests and have not been altogether satisfied with their results when they have engaged in test construction. On occasion, teachers have also been suspicious of standardized, commercial tests when they have not been sure what these tests were actually trying to measure.

With regard to *standardized achievement tests*, the national final exam which is annually conducted in Indonesia to evaluate the attainment of the National Education Standard (The Regulation of the Ministry of Education and Culture No. 66 Year 2013) was in fact turned into a method to assess the students' achievement for passing certain level of education. It created cumulative negative effects, as the students—especially the low achievers—became increasingly suspicious and cynical about them. Driana (2013) questioning the effectiveness of the national final exam for some reasons. A large number of students reported becoming anxious about such tests, cheating, trying halfheartedly, or using poor test-taking strategies such as not monitoring their time, not narrowing down the acceptable alternatives in multiple choice items, and not checking answer. She also emphasized that the national final exam could increase drop out students, narrow the curriculum that is focusing the instruction only on the subjects included in the national final exam (especially for ELT, less attention is provided for productive skills and traditional assessment is dominant), a lot of exercises (working on students' worksheet) must be accomplished by the students that limit the students' creativity, and more and more students get depressed. These attitudes and behaviors undermined the validity of the test scores and discouraged genuine learning. Therefore, developmental approach of assessment, whereby teachers and students would work collaboratively on authentic tasks which are longitudinal and multidimensional is recommended, and such kind of assessment procedure could be found in some alternatives in assessment.

Alternative assessment requires students to perform real-world tasks (authentic) that demonstrate meaningful application of essential knowledge and skills (Brown, 2004:252). In contrary with the standardized test, alternative assessment tend to be open-ended in its time orientation and format, contextualized to a curriculum, referenced to the criteria (objectives) of that curriculum, and likely to build intrinsic motivation. This contradictory characteristic results negative correlation between standardized test and

alternative assessment. As a technique increases in its washback and authenticity, its practicality and reliability tend to be lower. Conversely, the greater the practicality and reliability, the less beneficial washback and authenticity could be achieved. Therefore, alternative assessment may represent meaningful instruction.

Based on the observation and interview conducted with the junior high school English teachers in Bali who participated in the teacher certification program 2014, it was found that all teachers love to teach. It is because they enjoy what they do that they waded through a medley of pre-service teacher education courses, conquered the challenges of student teaching, and hopped the myriad hurdles of the certification process. Although the teachers like to *teach*, they rarely like to *assess* students' learning, especially by using alternative assessment even though knowing about assessment is really important to support the teachers' pedagogy competence. Popham (2011:8-23) proposes four traditional reasons why teachers must become knowledgeable regarding educational assessment, namely: (1) diagnose students' strengths and weaknesses,(2) monitor student's progress, (3) assign grades, and (4) determine a teacher's own instructional effectiveness. Three more current reasons teachers need to know about assessment are identified to be used to (1) influence public perceptions of educational effectiveness, (2) help evaluate teachers, and (3) clarify teacher's instructional intentions. It was emphasized that teacher should use the results of assessment to make better decisions.

The teachers' reluctance in conducting alternative assessment is due to some reasons. The primary reason was related to the implementation of national final exam which focuses on the receptive skills: reading and listening. Therefore, teachers emphasized their teaching to the content area instead of on how the students perform their ability in using the language in productive skills: writing and speaking. Therefore, tests dominate the assessment process to know the students' quality. As the result, students' often find it difficult to use the target language spontaneously and communicatively. Moreover, the process of how students' come to their decision in answering the exercises was not the main concern as long as the students' answer is correct. Most of the teachers consider the mid-test and final exam as the only source of data to record the students' achievement, therefore motivation and self-confidence also seems to be neglected during the instruction. In addition, big number of students in a class led the teachers' decision to use test, even though much time spent to construct it, but the teachers believe that test is easy to administer and score.

The other reason stated by the teachers was due to the implementation of Curriculum 2013 where the Regulation of the Ministry of Education and Culture No. 66 Year 2013 about Evaluation Standard is issued as a reference for the educators, schools, and government at elementary and high school level to plan, administer and report the result of the assessment. It is clearly stated that the students must be assessed due to their competence, therefore, this document provides clear picture about how to assess the students' competence based on certain domain: cognitive, affective and psychomotor. Unfortunately, the teachers found themselves confounded by the decision in choosing assessment

technique for assessing their students' competence and how to administer it. This confusion may inhibit the teachers' success in conducting the instruction.

In short, it is importance for the teachers to know some philosophical nature of assessment and alternatives in assessment, such as the reason why language ability must be assessed and what principles of language assessment are; and its practical side such as the implementation of alternative assessment techniques (portfolio for psychomotor domain, journal and self and peer-assessment for affective domain) based on the Evaluation Standard issued by the Indonesian Ministry of Education and Culture, the challenge, and opportunity in administering the assessment procedure. Thus, by having more insight about language assessment, especially alternatives in assessment, the teachers' doubtless could be reduced, as the result successful instruction cycle could be achieved.

WHY ASSESS LANGUAGE ABILITY?

Assessing language ability provides benefit not only for the learners but also for the teachers (Cohen, 1994:13-6). Conducting language assessment in the classroom promotes meaningful involvement of students with material that is central to the teaching objectives of a given course. For this meaningful involvement to take place, the goals of the assessment tasks need to reflect the goal of the course, and these goals need to be made clear to the students. Regular assessment of learning can provide learners with feedback about their language performance at various stages in the developmental process. While assessment is taking place, students are getting feedback on how well they perform on the tasks. Students may learn something about their areas of strength and weaknesses from the quality of the assessment and attention given to the students, prompting further learning or review.

For the teacher, assessing the students' language ability would provide the teachers information about how well the students did on the material being assessed and check for any discrepancies between expectations and actual performance. This information may indicate how well the students are learning if they have mastered the material, and how well the teacher has put across the material. Such feedback to the teacher can suggest areas for instruction, for review, or for improving future assessment.

PRINCIPLES OF LANGUAGE ASSESSMENT

Generally, language assessment should meet five principles according to Brown and Abeywickrama (2010:25-39), namely practicality, reliability, validity, authenticity, and washback. Practicality refers to the costs, the amount of time it takes to construct and to administer, ease of scoring, and ease of interpreting/reporting the results. Reliability means the consistency and dependability of the assessment instrument. Validity is the extent to which inferences made from assessment results are appropriate, meaningful and useful in terms of the purpose of assessment. Authenticity offers tasks that replicable real-world tasks, provides some thematic organization, such as through a story line or episode, and contains language that is as natural as possible. Washback refers to the effect of the assessment implementation which should be beneficial

for giving learners feedback that enhance their language development, contributing to the improvement of what and how learners learn and teachers teach. Due to the implementation of 2013 curriculum, the language assessment should follow the principles stated in the Regulation of the Ministry of Education and Culture No. 66 Year 2013 about Evaluation Standard. The assessment conducted for assessing the elementary and high school students must be objective, integrated, economical, transparent, accountable, and educative. Objective means that the assessment must be based on the standard and free from subjectivity. Integrated refers to a well-planned assessment that is conducted integratedly and sustainably within the instruction. Meanwhile, economical means the effectiveness of the assessment planning, administering, and reporting procedure. Assessment must also be transparent which means that the assessment procedure, criteria, and the basis of making decision may be accessed by the stakeholders so the accountability of the result could be guarantee. Finally, the result could motivate both students and teachers.

ALTERNATIVES IN ASSESSMENT: TECHNIQUES, CHALLENGES AND OPPORTUNITIES

In the public eye, tests have acquired an aura of infallibility. Everyone wants a test for everything, especially if a test is cheap, quickly administered, and scored instantaneously. However, test has its disturbing thing to an extent that many people accept the result uncritically, while others believe that all testing is invidious (Brown&Abeywickrama,2010:122). Therefore, alternatives in assessment is needed to be proposed to provide more fairness in assessing the students' language proficiency.

Portfolios

A portfolio is a systematic collection of students' work which display their skills and accomplishment (Popham, 2011: 211). Portfolio can be process, product and showcase portfolio. Johnson and Johnson (2002:106) provides some reasons why portfolio is used as an alternative assessment. Portfolio gives students the opportunity to direct their own learning. Students can document their efforts, achievements, development, and growth in knowledge, skills, expressions, and attitude. Students can also use various learning styles and intelligences, and they also have more chance to assess their own learning and decide which items best represent their achievement and growth so that they can set their future goal, so that the students' level of achievement could be determined. Moreover, portfolio can also be used to evaluate and improve curriculum and instruction.

A portfolio consists of some items. It includes cover sheet that creatively reflects the nature of the students (or group work), table of contents that consist of titles of each work sample and its page number, the rationale which explains what work samples are included, why each one is significant, and how they all fit together in a holistic view, the works samples, a self-assessment written by the students or the group members, future goals based on the students' (or group's) current achievement, interest and progress, and other's comments and

assessment from the teacher, cooperative learning group and other interested parties such as the parents.

How portfolio can be implemented? First, in the preparation step, decide what type of portfolio will be used. Then, identify the purposes and objectives of the portfolio. Next, select which categories of work samples should go into the portfolio. After that, have the students' select the pieces to be included in their portfolio. Finally, decide how the portfolio will be assessed and evaluated. As the portfolio is implemented, there will be a management step such as explaining the categories of work samples to be included in the portfolio, developing scoring rubric, completing the assignments by the students, and conducting self-assessment. The last step is conducting grading period. Once all the work samples have been completed, the selections for the portfolio must be made and organized into a coherent representation of the students' or groups' work.

In spite of its plus side (the potential to create authentic portrait of what students learn, the use of self-assessment guides the students' learning, and improving personal ownership), there are some challenges that teachers might face during the implementation of portfolio assessment as described by Popham (2011: 220-1). Sometimes the scoring guides used in evaluating portfolio are so general as to be almost useless. In contrast, some scoring guides are so detailed and complicated that they simply overwhelm scorers. Davis and Ponnanperuma (2005) as cited in Lotfi (2012) assert that difficulties may arise in verifying whether the material submitted is the learners' own work. Generally speaking, most teachers are so busy they don't have time to create elaborate scoring scheme. Another challenge is it takes time—loads of time—to carry out properly. Many hours should be devoted both inside and outside of classroom. To face these challenges, if teachers in school will be using portfolio assessment in their classes, this will be a marvelous opportunity to establish a *teacher learning community* in which, on a continuing basis during the year, portfolio-using teachers meet to share insights and work collaboratively on common problems.

Journals

A journal is a log of one's thoughts, feelings, reactions, assessments, ideas, or progress toward goals, usually written with little attention to structure, form, or correctness (Brown&Abeywickrama, 2010:134). Most classroom-oriented journals are what have now come to be known as dialogue journals. They imply an interaction between a reader (the teacher) and the student through dialogues or responses. Through this, teachers can become better acquainted with their students to meet student's individual needs. Tuan (2010) emphasizes that the journal writing could also foster writing motivation and enhance their writing skill.

There are some steps that could be followed in implementing journal for classroom instruction. First, students should be sensitively introduced to the concept of journal writing such as by providing examples of journal entries, given specific topics and schedule for writing. Then, state the objective(s) of the journal, such as for language-learning logs, grammar journal, responses to readings, strategies-based learning logs, self-assessment reflections, diaries of attitudes, feelings and other affective factors, and acculturation logs. After that,

give guidelines on what kinds of topic include and carefully specify the criteria for assessing or grading journals. Optimal feedback in teachers' responses should be provided, it could be in the form of cheerleading feedback, instructional feedback, or reality-check feedback. The step is continued to designate appropriate time frames and schedules for review, and finally provide formative, washback-giving final comments. Credit might be given for the process of actually writing the journal, and possibly a distinction might be made among high, moderate, and low effort and/or quality, but to accomplish the goal of positive washback, narrative summary comments and suggestions are clearly order.

With the widespread availability of internet communication, journals and other student-teacher dialogues have taken on a new dimension. With such innovations as "collaboratories", on-line education, and distance learning, journals--out of several genres of possible writing—have gained additional prominence. Because most journals are-or should be- a dialogue between student and teacher, they afford a unique opportunity for a teacher to offer a various feedback. On the other side of the issue, it is argued that journals are too free to be assessed accurately. With such potential variability, the concept of free and unfettered writing is anathema. Students may be asked to reveal an inner self, an act that may be virtually unheard of in their own culture. To answer the challenge, the reliability of the assessment can be maintained by the adherence of journal entries to stated purposes and objectives, but because individual variations in writing and the accompanying variety of responses, reliability may reach only a moderate level. Content validity and impact are very high if the journal entries are closely interwoven with curriculum goals (which in turn reflect real-world needs).

Self-and Peer Assessment
Self-assessment derives its theoretical justifications from a number of well- established principles of second language acquisition. The principle of autonomy stands out as one of the primary foundation stones of successful learning. The ability to set one's own goals both within and beyond the structure of classroom curriculum and independently monitor all keys to success. Peer-assessment appeals to similar principles, the most obvious of which is cooperative learning. The benefit of a community of learners which capable of teaching each other something. Peer-assessment is simply one arm of a plethora of tasks and procedure within the domain of learner-centered and collaborative education.

There are some types of self and peer-assessment, namely direct assessment of performance, indirect assessment of performance, metacognitive assessment, assessment of socio-affective factors, and student self-generated tests.

Self and peer-assessment must be carefully designed and administered to reach their potential. Four guidelines will help teachers bring intrinsically motivating task into the classroom successfully. First, tell students the purpose of the assessment, then define the task(s) clearly. After that encourage impartial

evaluation of performance or ability, and finally ensure beneficial washback through follow up tasks.

There are some possibilities for self and peer-assessment that could be considered as a taxonomy of self and peer-assessment tasks. For listening tasks, students may be asked to listen to TV or radio broadcast and checking comprehension with partners or asking when one does not understand something in pair or group. For speaking tasks, students may fill out student self-checklist and questionnaire, rating someone's presentation, or asking others for confirmation checks in conversational settings. For reading tasks, students may be asked to conduct self-assessment on reading habit, reading and checking comprehension with a partner, or setting goals for creating/increasing opportunities for reading. For writing tasks, students may be assigned to revise written work by their own, peer-editing, proofreading, or using journal for reflection.

Self and peer-assessment offer certain benefit. Brown and Hudson (1998) as cited in Brown and Abeywickrama (2010:145) clarify that self and peer-assessment improve direct involvement of students in their own density, encourage autonomy, and increase motivation because of the engagement. Of course, some noteworthy drawbacks must also be taken into account. Subjectivity is a primary obstacle to overcome. Students may be either to harsh on themselves or too self-flattering, or they may not have the necessary tools to make anaccurate assessment. Especially in the case of direct assessment of performance, students may not be able to discern their own errors. Moreover, students' local culture may also influence the students' ability in conducting self and peer assessment. For example, there is *de ngaden awak bisa* (do not think oneself knowing everything) concept in Balinese culture that affect the students' ability in making assessment (Marhaeni, et.al. 2005). This concept makes the students reluctant to judge themselves having good work, and at the same time they tend to judge their pair's work subjectively. But there is also *jengah* (being motivated) concept which may give positive effect toward the students' self-motivation. This is the teacher's task to overcome the challenge. Nevertheless, successful learners' extend the learning process well beyond the classroom and the presence of a teacher or tutor autonomously mastering the art of self-assessment. In addition, Yang&Tsai (2010) as cited in Kearns (2012) states that in order to reducing teachers' workload, rubrics can be as simple as a checklist that specifies target performance criteria for an assignment can be used to guide student activity. And where peers are available to render assessments, why not take advantage of such additional input?

CONCLUSION

2013 Curriculum implemented in Indonesia education system affects the assessment procedures applied for the instruction. Test is no longer becomes the only technique for assessing the students, more authentic and meaningful assessment must be conducted to assess the students' competence. There are some alternatives in assessment that can be selected, and for the purpose of this paper, portfolio, journal, and self and peer assessment are proposed to answer the teachers' problem in deciding appropriate assessment technique for assessing

the students. There are some challenges that might be faced by the teachers in implementing the assessment, such as the teachers' competence, teachers' workload, teachers and students time management, students' writing ability, students' learning autonomy, students' low achievement, and local culture influence. Therefore, to answer the challenges, there are some opportunities that could be considered, such as conducting community learning, integrating technology into the instruction, and accommodating local culture which gives positive washback effect for the assessment instrument development.

REFERENCES

Brown, H. D., 2004. Language Assessment. Principles and Classroom Practices. New York: Pearson Education

Brown H. D.,& Abewicakrama P., 2010. Language Assessement: Principles and Classroom
Practices. New York: Pearson Education

Cohen A. D., 1994. Assessing Langugae Ability in the Classroom. Boston: Heinle&Heinle

Driana. 2013. Masih Perlukah Ujian Nasional? Available at http://edukasi.kompas.com.
Accessed on 3 October 2014

Johnson, D. W. & Johnson, R. T., 2002. Meaningful Assessment. A Manageable and Cooperative Process. Boston: Pearson Education

Kearns, L. R., 2012. Student Assessment in Online Learning: Challanges and effective Practices. MERLOT Journal of Online Learning and Teaching Vo;. 8 No 3, Sepetember 2012. Available at http://www.jolt.merlot.org. Accessed on 5 October 2014

Lotfi H. A., 2012. Using Portfolio Assessment Technique in EFL Classroom: Necessity, Reduncy or Luxury. ELT Voices-India, (Online), 2 (6):93-103, (http://eltvoices.in/EVI26/EVI_26_6.pdf. Accessed on 5 October 2014

Marhaeni, Ramendra & Suwastini. 2005. Optimalisasi Kegiatan Evaluasi Diri Dalam Pembelajaran Berbasis Portofolio untuk Meningkatkan Kemampuan menulis dalam Bahasa Inggris. Singaraja. IKIP N Singaraja.

Peraturan Menteri Pendidikan dan Kebudayaan No. 66 Tahun 2013 tentang Standar Penilaian

Popham W. J., 2011. Classroom Assessment (Sixth Edition). Boston: Pearson Education

Tognolini, J & Stanley, G. (Ed.). 2011. A Standards Perspective on the Relationship between Formative and Summative Assessment. East Asia: Brithis Council

Tuan L. T., 2010. Enhancing EFL Learners' Writing Skill via Journal Writing. English Language Teaching Journal, 3 (3). (Online), (http://www.ccsenet.org/journal/index.php/elt/article/view/7217) Accessed on 5 October 2014