MULTIPLE ASSESSMENT METHODS:
TOWARDS THE IMPROVEMENT OF STUDENTS' ENGLISH LANGUAGE LEARNING

Sugiono
The University of Nurul Jadid
ss.sugiono@gmail.com

Abstract

The study explores the perspective of multiple assessment methods and possibility of applying it in English language learning. Assessment of English language learning needs to consider the philosophical perspective of progressive education. Assessment in progressive notions of education is primarily used to nurture and develop individual students' capacity and range of talents. Likewise, the perspective of the 2013 National Curriculum demands students' real performance on relevant tasks. The study notes that an assessment process refers to implementing multiple methods or a range of strategies to make decisions regarding instruction and gathering information about students' performance and behaviour, diagnoses students' problems, monitors their progress and gives feedback for improvement. Rather than a single method of assessment, multiple assessment methods are more beneficial to help the teacher triangulate the evidence for a complete picture of student comprehension and understanding. Every method of assessment has weaknesses, and hence, by using multiple assessment methods, the strengths of one particular method can overcome the problematic weaknesses of another.

Keywords: multiple, assessment, CAT, improvement

INTRODUCTION

Since mid-1998, the locus of Indonesian educational decision making has shifted from a more centralized, bureaucratic, authoritarian model to a decentralized, democratic, inclusive mode of operation. Suggested reforms for organizational systems have been accompanied by efforts to introduce reform in curriculum, instruction, and evaluation. One of the results is the design of the national curriculum – the new 2013 Curriculum – which has offered the promotion of schools' and teachers' autonomy in the provision of educational practices and assessment. By referring to the broad guidelines of the national curriculum, schools have an authority to set up their own assessment method known as Classroom Assessment Tasks (CAT).

CAT is an activity of gathering information about students' learning outcomes and processes so as to measure what is intended to measure. One of the basic principles of CAT is that it is developed by teachers as teachers know more about their students' learning achievements. They responsible for the development of students' learning in that they are front-line problem-solvers of schooling and educators having more knowledge of students' needs (Print 1987, p. 15). "Researchers estimate that teachers may spend from one third to one half of their time in assessment-related activities" (Stiggins, Conklin, & Associates 1992 as cited in Nitko & Brookhart 2007, p. 3).

Other fundamental principles of CAT are: Inseparable from the teaching-learning process, using criterion-referenced measures, using various methods of assessment (test and non-test), orienting itself to students' comprehensible competence, valid, fair, open, interactive, on-going, meaningful, and educative. Hence, CAT can include portfolios, products, projects, performance, and written tests (paper-and-pencil tests).

With reference to the aforementioned principles and coverage of CAT recommended by the national curriculum, it could be stated that in general English language learning needs to submit to the philosophical perspective of progressive education. Assessment is one essential part of progressive notions of education, which are mainly functioned to nurture and develop individual students to their full capacity and use their complete range of talents (Brady & Kennedy 2005). Likewise, the perspective of the new curriculum demands students' real performance on relevant tasks.

MULTIPLE ASSESSMENT METHODS IN ENGLISH LANGUAGE LEARNING

According to Butler and McMunn (2006, p. 2), an assessment process refers to using multiple methods or a range of strategies to make decisions regarding instruction and gathering information about students' performance and behaviour in order to diagnose students' problems, monitor their progress, or give feedback for improvement. Using multiple and diverse sources, teachers can be sure that they have a more complete and accurate view of each student's level of achievement in relation to the learning targets.

The philosophical framework of the 2013 National Curriculum, and the development of assessment system, recommend seven assessment methods selected to be adapted to the sphere of English language learning. These are Simple Product Assessment, Performance Assessment, Logs, Journals, Notebooks, Portfolios, and Project Assessment.
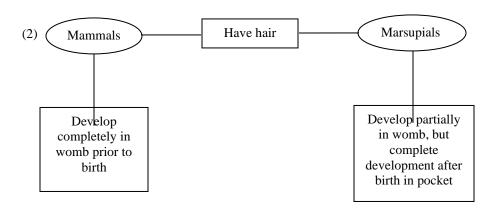
Simple Product Assessment

This assessment of students' English language learning can include short answer sentences or paragraphs, diagrams and illustrations, graphic organizers, graphs and tables, and matrices. Short-answer questions or paragraphs are used to answer particular questions, which recalls student-constructed responses. According to Popham (1978, p. 64), "the primary weakness of short-answer items relate to the difficulty in scoring them satisfactorily, for it is possible that the student will construct a variety of responses".

Apart from that, there is the problem of penmanship, where some students may produce a correct answer which is literally indecipherable.

Despite these drawbacks, as students do not select the correct answers from the list but construct the response, there is a good chance for them to enhance their understanding or comprehension of particular topics. Besides, like short-answer questions or student-constructed paragraphs, diagrams and illustrations, graphic organizers, graphs and tables, and matrices are other examples of student-constructed responses but use pictorial displays to uncover student knowledge. For examples,

(1) There are four wooden boards, each 16 feet long. If you used these boards to lay out a square on a playground for a proposed sandbox, how large an area would the square encompass?

(2)

```
  ( Mammals )────[ Have hair ]────( Marsupials )
      │                                 │
  ┌─────────┐                   ┌──────────────────┐
  │ Develop │                   │ Develop partially│
  │completely│                  │  in womb, but    │
  │in womb   │                  │    complete      │
  │prior to  │                  │development after │
  │  birth   │                  │ birth in pocket  │
  └─────────┘                   └──────────────────┘
```

Sample (1), *short-answer sentences*, can be used to check comprehension or to explore the recall of information by students, while sample (2), *graphic organizers,* is for representing key skills like sequencing, comparing and contrasting, classifying and involving students in active thinking (Burke 1994, p.118 as cited in Butler and McMunn 2006).

Performance Assessment

In English language learning, the performance assessment is part of multiple intelligences approach (Butler and McMunn 2006, p. 254), which is concerned with skill outcomes. In performances, an audience (=teacher) is usually present to watch the students. Like products, performances need to conform to the definition of constructed response assessments in that they must be judged according to pre-established performance criteria. Like all quality assessments, performances should also help forward the instruction of key curricular concepts (Butler and McMunn 2006, p. 59).

Performance assessment can be applied to various disciplines. While performances in science and mathematics courses are concerned with laboratory and practical problem-solving skills, those in English language learning are dealt with communication skills. In addition, performances can be presentations, demonstrations, enactments, debates, panel discussions, and videotapes or audiotapes. Thus, typical action verbs used for performance outcomes are 'construct', identify', and 'demonstrate'. In this case, performance assessment places greater emphasis on the application of

knowledge and skill, which results in desired learning outcomes. It is also used as an instructional device to prepare students for actual performance in real situations and as evidence of individual students' ability to apply their English language skills in an actual context.

Performance assessment is frequently neglected in the measurement of instructional outcomes in that it is more difficult to use, requires more time to prepare and administer, and scoring it frequently tends to be subjective and burdensome (Gronlund, 1990, p. 81). However, Popham (1995, p. 139-153) suggests that significant instructional improvement will flow from using such an assessment, for it represents an alternative to traditional paper-and-pencil test, and is often more authentic and reflective of tasks that people need to perform in the real world. Another big plus of performance assessment is that it establishes assessment targets which, as the targets often influence teacher's instruction, have a positive impact on instructional activities. Using performance assessment is also believed to be able to raise standards as long as it requires students to complete a well-crafted performance task, providing them with the opportunity to apply their English language learning to a new situation and shares teachers' scoring rubrics with students to clarify the learning targets for them (Butler and McMunn 2006, p. 253).

Logs, Journals, and Notebooks

Logs, journals, and notebooks constitute the types of written documents that students produce and are used as assessment evidence. *A log* provides documentary evidence of events and concise summaries of information, which show the progression of events. English language teachers employ this assessment method when students need to track information and use it for later assignments. In schools, logs are used to verify students' actions. A detailed log helps convince the teacher that the students perform certain actions. Moreover, the log can reveal the exact nature of these actions. Because of their documentary properties, logs are frequently used to support assertions or conclusions students make about information they have gathered.

Journals are similar to logs, that is, to provide a record of the progression of events. In a journal, students record thoughts, observations, and questions and hold written dialogues with others. Nevertheless, journals do not generally have the legalistic and evidentiary purpose of a log. English language teachers can select this method to require students to record specific information based on the learning targets. A diary is an example of journals. Even though the diary contains personal events, it flavours these events with the opinions, feelings, and perceptions of the owner. It is this flavouring which makes journals more useful to teachers to get information on the improved English language learning of students.

Assessment which uses *notebooks* as students' written products has been widely implemented in educational settings from primary to secondary levels. A notebook is similar to a file folder in which it commonly holds a collection of all information pertinent to a particular topic. Students create their own 'notebooks' at schools by voluntarily compiling notebooks for their classes. An English language notebook, for instance, can contain notes during learning activities, copies of exams, and completed

homework assignments. The students collect this information by means of the notebook as a study guide for upcoming exams. The completeness of information is of primary importance in order for them to be successful in English language learning. In other words, notebooks are helpful for assessment purposes, for they contain the totality of work produced by students. The data as such can be analysed to track student performance over time, determine particular content areas or concepts in which the students experience difficulty, or serve as a basis of self-assessment.

Portfolios

In education, portfolios refer to systematic collections of students' work. Unlike notebooks, they do not usually include the totality of student's work. A portfolio is purposeful, integrated collection of student work showing effort, progress, or a degree of English proficiency. Most advocates of portfolio assessments believe that the real payoffs for such assessment approaches lie in an individual teacher's classroom. The relationship between instruction and assessment is strengthened due to students' continuing accumulation of work products in their portfolios. The English language teachers who adopt portfolios in their classrooms potentially make the on-going collection and appraisal of students' work a central focus of the instructional program rather than a peripheral activity whereby students occasionally gather up their work to convince teacher's supervisors or students' parents that good things have been going on in the classroom.

Popham (1995, p. 171), however, specifies the downside of portfolio assessment, stating that portfolio assessment commonly suffers from the limitation of all constructed-response measurement. Students' constructed responses are genuinely difficult to evaluate, particularly when these responses vary among students. Another problem with portfolio assessment is that it takes time to carry out properly. Being efficient in reviewing students' portfolios, teachers still have to devote many hours both in class (during portfolio conferences) and outside the class (if also wishing to review students' portfolios by themselves).

Despite the aforementioned disadvantages, portfolio assessment provides advantages, among others are, engage students in assessing their progress and/or accomplishments and establishing learning goals, measure each student's achievement while allowing for individual differences between students, have a goal of students' self-assessment, and link assessment and teaching to learning.

Implementing portfolios to improve learning, teachers need to be aware of portfolio assessment specialists suggest, that the most important dividend from portfolio assessment is the increased abilities of students to evaluate their own work. If this is one of the goals, teachers need to be certain to nurture such self-evaluation to grow deliberately through portfolios instead of simply using portfolios merely as convenient collections of work samples to appraise.

In addition, to enhance the validity and reliability of portfolio assessment, the materials, resources, and procedures need to correspond to those of instruction. Students need to understand the goal, content and evaluation criteria of their portfolios. Using the

portfolio assessment in the classroom, the teacher needs to be consistent with the criteria towards the whole class, and students ought to participate in the development of rubrics along with their teacher. It is contended that reaching close unanimous agreement among teachers on ratings within a rubric on individual entries or the portfolio as a whole potentially provides consistency or reliability.

Project Assessment

How projects are used in assessment? Like portfolios, projects are powerful assessment methods. However, what makes projects different from portfolios is that projects do not encourage students to house a collection of work but rather demonstrate mastery of skills or completion of tasks (Butler & McMunn 2006, p. 74). In this respect, projects appear under both the product and performance classifications. To show such mastery, students need to create written products and perform demonstrations. Besides, projects can offer students an opportunity to direct their own learning and integrate content across curricular.

In other words, projects need to be designed and selected to teach core curriculum content standards and scored by using a rubric, which is shared with students in advance. To support student autonomy and decision making, students are given some choices as to the tasks they will perform or the roles they will assume for the project. Students are required to meet interim deadlines for the project, participate in planning the project, and reflect on project activities. In each case, both the process and product are assessed. Students ought to understand the criteria used for both. For instance, some processes include use of class time, resources, engagement, and time management, whereas the actual product criteria incorporate content knowledge, method of presentation, and accuracy of the information.

The primary limitation of a project, however, refers to assessor's involvement in the project and the research process. In this case, Shay (2005, p. 670) considers that this involvement can impact on teacher's objectivity in marking the project, for, spending several months agonizing with students, the teacher knows how many of the ideas are theirs. Nonetheless, this drawback can be overcome by providing two or three assessors of the project. By so doing, analysis and synthesis of data from multiple sources can be developed to reduce the subjectivity of the single assessor.

Concluding Remarks: Socializing Innovative Multiple Assessment Methods

To the educational setting in the institution or schools in Indonesia, socializing and introducing such assessment methods is a challenging thing. A number of local educational experts note several problems that can come up from the socialization process:

(1) *Lack of time for English teachers to undertake multiple assessment methods.* English teachers are burdened with a heavy load of teaching hours, which is up to 42 hours a week. Their income depends on the amount of their teaching hours. The more teaching hours they have, the more income they earn.

95

(2)   Lack of English teachers experienced or trained in the kind of assessment. It is contended that such assessment methods require significant changes in the roles of teachers that have resisted changing due to a long engagement in the old system of education.

(3)   Large classes. With 35 to 45 students in one class makes English teachers difficult to effectively organize or manage the class.

Noting that almost all schools are closely connected with the state's policies and "most schools are government run" (Jewell 2005), there should be government's involvement in introducing the innovative multiple assessment methods, in the form of either providing intensive trainings on applying multiple assessment methods or regularly supervising the implementation. Then, the major things that should be firstly explained to the stakeholders are that assessment is inseparable from the curriculum development and integral part of the educational process, which keep on providing feedback to improve students' learning. Besides, promotion of students' learning is a major aim of schooling, and assessment is at the centre of the process and needs to be integrated into teaching strategies and practices at all levels.

In addition, making use of multiple assessment methods will provide more benefits of helping the teacher triangulate the evidence for a complete picture of student comprehension, compared to using a single method of assessment. Another reason is that every method of assessment has limitations or weaknesses. Accordingly, by using multiple assessment methods, the strengths of one particular method, for instance, can overcome the problematic weaknesses of another.

REFERENCES

Alderson, JC 2000, 'Technology in testing: the present and the future', *System 28,* Accepted 10 February 2000, <www.elsevier.com/locate/system>.

Brady, L & Kennedy, K 2005, Celebrating student achievement: Assessment and reporting, Prentice Hall, Frenchs Forest, N.S.W.

Butler, SM & McMunn, ND (2006). A Teacher's Guide to Classroom Assessment: Understanding and Using Assessment to Improve Student Learning. San Francisco: John Wiley & Sons.

Gronlund, NE 1990, Constructing Achievement Tests, Englewood Cliffs, N.J.

Henning, G 1990, A Guide to Language Testing, Newsbury House Publishers, Cambridge.

How standardized testing damages education, n.d., the National Centre for Fair & Open Testing, Cambridge, <www.FairTest.org>.

Jewell, P 2005, 'Autonomy and Liberalism in a Multicultural Society', International Education Journal, September 2005.

Linn, RL & Gronlund, NE 1995, Measurement and Assessment in Teaching, Merrill, Englewood Cliff, N.J.

Nitko, AJ & Brookhart, S 2007, Educational Assessment of Education, Pearson Education, Australia.

Popham, WJ 1978, Criterion-Referenced Measurement, Englewod Cliffs, New Jersey.

Popham, WJ 1995, Classroom Assessment: What Teachers Need to Know, Allyn and Bacon, Boston.

Print, M 1987, Curriculum Development and Design, Allen & Unwin, Australia.

Shay, S 2005, 'The assessment of complex tasks: a double reading', Studies in Higher Education, Vol. 30, No. 6, December 2005, pp. 663-679.

Wanted: Creative Teachers, 2006, Tempo, March 20.

Wiggins, G n.d., The Case for Authentic Assessment, Viewed 29 April 2006, <http://www.educationoasis.com/curriculum/assessment/caseauthentic assessment.htm>.