
AUTHENTIC ASSESSMENT IN ELT: HOPES, CHALLENGES, AND PRACTICES

Tono Suwartono¹ and Cici Riyani

¹Universitas Muhammadiyah Purwokerto, ²SMA Negeri 1 Baturraden

Email: suwartononewton@gmail.com; ciciriyani77@gmail.com

Info Artikel

Sejarah Artikel:

Diterima 2 Januari 2019

Direvisi 8 Maret 2019

Disetujui 10 Mei 2019

Keywords:

Authentic assessment, English Language Teaching (ELT), hopes, challenges, practices

Abstract

The aim of this paper is to provide an overview of the need of authentic assessment in English language teaching in Indonesia. After the implementation of 2013 curriculum (Discovery Learning Curriculum), the evaluation that must be done by teacher is not only emphasized on cognitive domain but also in psychomotoric and affective domains. This is a critical idea based-article using a literature review method. The authors reviewed the authentic assessment from both the information literacy and instructional design which can be applied by EFL teachers in their classes. Overall, this paper gives an outlook on what, why and how all about authentic assessment in ELT.

Abstrak

Tujuan artikel ini adalah untuk memberikan gambaran tentang perlunya penilaian otentik dalam pengajaran bahasa Inggris di Indonesia. Setelah implementasi kurikulum 2013, evaluasi yang harus dilakukan oleh guru tidak hanya ditekankan pada ranah kognitif tetapi juga pada ranah psikomotor dan afektif. Artikel ini berdasar pada ide kritis yang menggunakan metode kajian pustaka. Penulis mengulas penilaian otentik dari literasi informasi dan desain pengajaran yang dapat diterapkan oleh guru di kelas. Secara umum, artikel ini memberikan pandangan tentang apa, mengapa, dan bagaimana penilaian otentik di dalam pengajaran bahasa Inggris.

INTRODUCTION

English Language Teaching (ELT) is a program. Main components of a program include goal, implementation, and evaluation. Any program needs evaluation. In its broader sense, evaluation is a process which determines the extent to which objectives have been achieved. In the education program, more specifically in the context of teaching, the evaluation methods could be test or non-test. However, before the so-called “2013 Curriculum” – which globally refers to Discovery Learning – was launched, evaluation had been always linked to tests and examination due to the emphasis of cognitive domain in the earlier KTSP (Education Unit Level Curriculum). It has been a revolutionary change from this “mere” cognitive domain evaluation mode to more holistic domains of cognitive – affective – psychomotoric evaluation method implemented in the recent curriculum.

AUTHENTIC ASSESSMENT: MORE THAN JUST A TEST

Tests methods could be objective tests as opposed to subjective tests, spoken tests as opposed to written tests, and so on. Non-test methods could be observation (using checklists, observation sheet), questionnaire (e.g. rating scales), journals (e.g. logs, diaries), document (e.g. portfolios), etc. So, time has changed. Now with the 2013 Curriculum, classroom practioners without exception, willy nilly, have to become familiar with techniques for data collection like researchers. It is not surprising, as a teacher serves also as a classroom researcher.

Some terms that are closely associated with evaluation in education program are assessment, test, examination, rating, marking, and scoring. Assessment, according to many articles in International Journal of Testing (IJT), covers data collection and decision making activities. Tests and examinations are quite similar (no need to distinguish!). Rating, marking and scoring are also similar. Since the 2013 Curriculum was implemented, the so-called Authentic Assessment has been popular. The rest of this paper will discuss assessment and authentic assessment in particular.

Let us start with assessment in relation to teacher's role. As touched on previously, assessment covers the activities of data collection and decision making. In the earlier curriculum, a teacher collected data on student competence, knowledge, performance, ability, skill, and mastery through tests for a decision. As a form of assessment, tests and testing are very crucial.

First of all, “...tests are important tools with significant consequences for the persons tested and organizations using the scores...” (Evers, *et. al.*, 2017). Secondly, tests and testing are as important as teaching. It is because tests and testing can bring effect to teaching and learning. Some writers use term “washback” or “backwash” for describing the effects or influences brought by tests or examinations (Pan, 2016). Numerous studies have shown that test and testing influences beneficially and harmfully in teaching and learning. If the effect is beneficial, then it is stated as positive backwash and if the effect is harmful, it is called as negative backwash. Siddiek (2010) states that positive backwash can yield good influence on both the teacher and learner and negative backwash can yield bad results on the teaching and learning processes. Negative backwash happens if the test material is not consistent to teaching content. Just imagine what happens to students if your test material is not consistent to your teaching content? Now, with the recent curriculum, a teacher needs to use methods other than test too for a better insight and right decision, i.e. the one that considers varied aspects. This refers to authentic assessment. To most teachers this dramatically leads to a harder job!

In the field of assessment, traditional assessment methods are often contrasted to authentic (real world) assessment. Other sources label traditional assessment as formal assessment, while authentic assessment is also known as informal assessment, performance-based assessment, or alternative assessment. However, those all are just names. Sunarti and Rahmawati (2014) and also Lampiran Permendikbud RI tentang Standar Penilaian Pendidikan (2013) define authentic assessment as an assessment that is done comprehensively to assess learning input, process, and output. By doing so, the assessment covers all domains of cognitive, affective, and psychomotoric as to consider learners characteristics and level. In the lower level affective aspect is loaded more heavily. In contrast to, in the higher level cognitive aspect is given more portions.

While in traditional assessment an assessment is very formal in the sense it is done at the end of teaching, in authentic assessment an assessment goes along the way the teaching, in such a way that both seem inseparable. In other words, assessment is done continuously and sustainably in the course of teaching and learning process. This seems realistic because in real life

situation learning takes place this way. It can be achieved by using varied evaluation methods. The advantage is that the teacher is able to obtain the whole picture of learners' learning progress, to monitor process, progress, and continuous improvement. With regard to domains coverage authentic assessment does not only measure what learners know, but also what they can do.

WHY IS IT THEN AUTHENTIC ASSESSMENT?

Long before the present curriculum was implemented, a teacher of English has been supposed to see that in communicative language teaching (CLT), the capacity or ability to use language communicatively involves both competence and demonstration of the ability to use this competence. Therefore, the performance tasks candidates are faced with in communicative tests should be representative of the type of task they might encounter in their own real life situation and should correspond to normal language use where an integration of communicative skills is required with little time to reflect on, or monitor language input and output. The emphasis is no longer on linguistic accuracy, but on the ability to function effectively through language in particular contexts of situation.

It has been unfair so far, for example, that students are not sufficiently trained to write just because teachers assume that writing is difficult both to teach and rate. While this could be a matter of ignorance in the part of teachers, writing can be taught and assessed in a simpler and more communicative way. With technology advancement, students' writing performance can be assessed more practically by using certain scoring programs (see Almond 2014).

Teachers can also administer more communicative tests of English such as: 1) Students write a letter to the editor in response to a newspaper editorial about a local issue, 2) Students discuss and analyse a current event, 3) Students write a journal entry after reading a thought-provoking story, 4) Students write and illustrate their own autobiographical memoirs, 5) Working in pairs, one student assumes the role of a character from a story read in class, while the other student interviews the character to learn about his or her life, 6) Students write and perform poetry to demonstrate understanding of figurative language.

Likewise, in discovery learning, where the teacher employs real-world teaching, assessment should be in favour of the teaching.

Let us say Project-Based Learning explicitly mentioned within the recent curriculum. Alternatively, Social Investigation or Social Inquiry where learners are involved in intensified team-work and student-centered learning mode, it might seem unfair if their mastery is measured using language only subskill test of grammar or vocabulary. Even, it remains unfair if their learning mastery is measured using isolated skill test of writing, speaking, listening, or reading.

As far as discovery learning is concerned, assessing students using traditional assessment modes (i.e. essays, multiple choice, fill-in-the-blank, etc.) will also not make sense. Simply testing an isolated skill or a retained fact does not effectively measure student's collective abilities. Limited English proficient students frequently have difficulty expressing their mastery of content unless they have a certain level of English proficiency. These students with collective abilities but inadequate English proficiency will take the benefit of more varied assessments, such as portfolios, mini-projects, products, and self-assessment.

AUTHENTIC ASSESSMENT IN ELT

As briefly discussed above, authentic assessment covers wider domains of cognitive, affective, and psychomotoric through more varied methods of tests and non-tests.

1. Assessing Cognitive Domain

According to Bloom's taxonomy, cognitive domain consists of six levels, from the simplest level to more complicated one, namely: knowledge, understanding, application, analysis, synthesis, and evaluation. These can be assessed through tests/questions or assignment.

Let us assume that good teachers of English recognise tests and testing, and therefore, it is unwise to discuss tests and testing here in more details.

a. Written test

1) Subjective Test (Essay, Composition, etc.)

Example:

Directions: Write a paragraph of about fifty-five words describing a store or business that you know very well. Base your paragraph on answers to the following questions.

1. What is it called?
2. When did it start to do business?
3. How many employees does it have?
4. What do the employees have to do?

5. Does it have a lot of customers/clients? Why (not)?

In order to help teachers assess testees' work a rubric or scoring criteria is required. Consider the following:

Mechanics	20%
Vocabulary choice and diction	20%
Grammar & usage	30%
Organization	30%
TOTAL	100%

- 2) Objective Test (Multiple-choice, Matching, etc.)

Sample (error identification):

Directions: Identify the one underlined word or phrase, (A), (B), (C), or (D) that should be corrected or rewritten.

One of the girls lost her wallet and

Consider the following (sample):

Accent	
1	Pronunciation frequently unintelligible
2	Frequent gross errors and a very heavy accent make understanding difficult, require frequent repetition
3	Foreign accent requires concentrated listening, and mispronunciation lead to occasional misunderstanding and apparent errors in grammar and vocabulary
4	Marked foreign accent and occasional mispronunciations which do not interfere with understanding
5	No conspicuous mispronunciation, but would not be taken for a native speaker
6	Native pronunciation, with no trace of foreign accent
Grammar	
1 – 6	
Vocabulary	
1 – 6	
Fluency	
1 – 6	

When a rubric itself is considered not enough, do scoring using an inter-raters procedure.

- c. Assignment/Projects

Example (Pronunciation Class):

Find a video on YouTube that shows a conversation in which wh-questions are used. Make a video clip of similar kind in which wh-questions are used and submit it to suwartono2006@yahoo.co.id no later than 1.00 PM, May 10, 2017.

A scoring rubric is also needed here. It might read as follows:

Fluency	25%
Intelligibility	25%
Grammar	15%
Vocabulary	15%
Focused material (?)	20%

A

doesn't have some money for the

B

C

D

game

- b. Spoken/oral test

The formats for the testing of oral test or techniques of elicitation are interview, interaction with peers, response to tape-recordings, pictures, role play, reading aloud (not recommended), interpreting, imitation (not too short sentences). Which one is chosen depends a lot on learner's level and scoring criteria.

Example (role play):

A friend invites you to a party in an evening when you want to stay at home and watch the last episode of a television serial. Thank the friend and refuse politely. (The friend is acted out by the tester)

2. Assessing Affective Domain

Non-test assessment methods can be used to measure affective domain. Affective domain covers perception, interest, attitude, emotion, and values. With regard to English as foreign language two very important variables are interest and attitudes. Questionnaire and observation are mostly popular for collecting data on these variables.

Attitude is a tendency to respond to an object, situation, concept, or person, either like it or not like it. Student attitude toward English

as a school subject constitutes an indicator of the teacher's success in conducting the English lesson. Many studies show the more positive student attitude is the more likely he/she will achieve success in learning. The most popular instrument for collecting data on student attitude is questionnaire. Here is a sample of attitude questionnaire.

No	Statements	Scale			
1	I like studying English	4	3	2	1
2	I like doing English assignments				
3	I like asking the English teacher about the lesson				
4	I like reading English books				
5				

Interest is learner's curiosity to an object condition organised through experience that drives someone to have the object, understanding, and skills for attention or achievement purpose. Research shows the higher student interest is the more likely the students will achieve success. Data on student interest can be measured by using questionnaire. The following is a sample of interest questionnaire.

No	Statements	Scale			
1	I pay attention to English lessons	1	2	3	4
2	The English lessons are useful for me				
3	I am, happy when the English teacher is teaching				
4				

To prepare a questionnaire, as any other instruments for data collection, first of all, we must think about conceptual definition (aspects to cover) and operational definition (indicators). One thing necessary to note is that one indicator cannot always be broken down to more than one item – under such condition then the indicator itself is the item (question/statement). In addition to questionnaire, data on affective domain can also be gathered through observation. Consider the following observational checklist of student activity in discussion:

No.	Statements	Category		
		good	fair	Poor
1.	Actively sharing ideas in front of friends			
2.	Actively asking questions			
3.	Actively giving response to friends' ideas			
4.	Actively completing group discussion task			
5.	Obeying rules in discussion			
6.	Being responsible of group task			
7.	Together doing group task with friends			
8.			

Other methods that can be taken to assess affective domain are journal reports, inter-peer assessment, and interviews.

3. Assessing Psychomotoric Domain

Assessing performance is done by observing the students doing something. The aspects assessed consist of: 1) imitating, 2) constructing, 3) doing with a procedure, 4) doing well and

precisely, and 5) doing an action naturally. In ELT, performance is language use, either spoken or written. In order to be practical when doing assessment a rating skill or checklist is needed. For instance, teacher asks students to deliver a speech. For assessing students' speech, teacher can use check list and rating scale as in table 1 and table 2 below.

Tabel 1.

Check list for assessing psychomotoric domain

1) Check list

Name :
Class :

No.	Aspect	Good (Score : 2)	Average (Score : 1)	Low (Score : 0)
1.	Organization (<i>Introduction, body, conclusion</i>)			
2.	Content (<i>depth of knowledge, logic</i>)			
3.	Fluency			
4.	Language:			
	<i>Pronunciation</i>			
	<i>Grammar</i>			
	<i>Vocabulary</i>			
5.	Performance (<i>eye contact, facial expression, gesture</i>)			
Score				
Total Score				

Tabel 2.

Rating scale for assessing psychomotoric domain

Name :
Class :

No.	Aspect	1	2	3	4
1.	Organization (<i>Introduction, body, conclusion</i>)				
2.	Content (<i>depth of knowledge, logic</i>)				
3.	Fluency				
4.	Language:				
	<i>Pronunciation</i>				
	<i>Grammar</i>				
	<i>Vocabulary</i>				
5.	Performance (<i>eye contact, facial expression, gesture</i>)				
Score					
Total Score					

Competency:

1 = low

2 = average

3 = good

4 = very good

a. Projects

Project based learning uses learning by doing principle (Sa'idah, Yulistianti, dan Farida, 2017). For assessing the learning, it is done by evaluating an assignment which covers the activities of designing, implementing and reporting. The assignment must be completed within a certain

period/time (Ditjen Pendidikan Dasar dan Menengah, 2017). For example, teacher gives an assignment that students in a group of 5 must create a video of narrative story.

For assessing students' work, teacher can use both check list and rating scale.

1) Check list

Group :
 Class :

No.	Aspect	Good (Score : 2)	Average (Score : 1)	Low (Score : 0)
1.	Designing	2019		
2.	Implementing			
3.	Reporting			
Score				
Total Score				

2) Rating scale

Group :
 Class :

No.	Aspect	1	2	3	4
1.	Designing				
2.	Implementing				
3.	Reporting				
Score					
Total Score					

Competency:

1 = low

2 = average

3 = good

4 = very good

b. Portfolios

Several studies (e.g. Hung, 2016) report that portfolio is good for use in language teaching and assessment. It is supported by Akhter (2017) that using portfolio in language teaching; knowledge can be shown by students and information to judge students in their learning progress can be accumulated by the teacher. Portfolio assessment alleviates students' anxiety towards English tests. Some students also expressed they have higher English learning motivation while their teachers implemented portfolio assessment to replace traditional assessment. This authentic assessment can directly portrait the students' attitude, knowledge

and skill is portfolio. It is an assessment based on a collection of reflective-integrative information that shows the development of students' abilities in a given period. Practically speaking, with portfolio students may collect their work in a folder and will submit it to the teacher at a specified time. For instance, a teacher will assess students writing skill whose topic is 'My Diary'. For this, students should write minimum two products of written form work every week for three months and collect them in a folder. At a specified time, students submit their works to the teacher and the teacher can use the form below to assess the students' works, for example:

No.	Student's name	Preparation	Implementation	Results	Report	Score	Mark
1.	Ahmad	60	75	75	70	280	70
2.	Budi	75	80	70	75	300	75
3.	Citra	50	60	65	60	235	59
4.	Deni	60	65	R365	65	255	64
5.	Endang	75	70	70	70	285	71

c. Products Assessment

Products assessment is an assessment of the students' productive skill. In learning English, the skills assessed are speaking, reading, listening, and writing. From the four skills, speaking and writing are skills which are

usually assessed by teachers as a product assessment. For assessing the students' work, teacher can use rubric and a teacher is also free to choose which rubric is appropriate for the need.

Types

1. Oral interviews

2. Story or text retelling
3. Writing samples -- generating a text, paper
4. Projects/exhibitions
5. Experiments/demonstrations
6. Constructed-response items --responding in writing to open-ended
7. Portfolios -- focused collection of student work to show progress over time

THE WEAKNESSES OF AUTHENTIC ASSESSMENT

As written above, authentic assessment to most teachers dramatically leads to a harder job. It is because applying the authentic assessment is time-consuming. Take for an example, a teacher who teaches 24 hours per week has to teach 12 classes each of which consists of 36 students. It will be equal to a 432-student work correction. It does not only take time and energy to do, but also there will be threat of evaluator bias (accordingly in need of inter-raters procedure and rubric).

Aliningsih and Sofwan (2015) state that the teachers' main problem in applying the authentic assessment is that they have large number of students, exhausting and time-consuming activities, and complicated administration. Therefore, it is challenging for teachers of English to implement authentic assessment.

THE BENEFITS OF AUTHENTIC ASSESSMENT

Good chance for students who have high anxiety level under traditional testing situations (heavily language dependent). Not only the language is important in learning a language, content is also important. What happens if a student writes in a very good English, but unfortunately it does not make sense? Empowering students to take control of their own learning/to become independent thinkers and users of language. Bridging the "distance" between learning within the walls of the school and that of informal life (no wonder if adults outside the schools and parents are involved).

Those all above are promising in the sense that our English language teaching, learning, and testing would most probably be better sooner or later with the implementation of authentic assessment. It is because authentic assessment is beneficial in most learning contexts and it is necessary for every age (Reynisdottir, 2016). It is also stated by Fox, et

al, 2017) that authentic assessment also touches a socio-cultural standpoint; it offers students the opportunity to learn through engagement peers and with the wider community. Complaint in the part of teachers is most likely a temporary reaction over a demanding new system, until they recognise this as a valuable investment that gives positive impact on their teaching and learning, as wise words say 'no pain no gain'.

CONCLUSION

In the context of teaching, authentic assessment uses multi-sources of data to get a comprehensive picture of student's capacity with regard to a particular course. The multi-sources of data demands that teachers have good teaching learning administration such as teaching journal, students' attitude assessment, students' knowledge assessment, students' psychomotoric assessment, teaching learning diary, etc. It is hoped that the administration make teachers stay comfortable in their class and feel more enthusiastic in teaching their students.

To better practice authentic assessment willy nilly a teacher must learn data collection (and, of course. data analysis) techniques. By doing so, an EFL teacher also practices their knowledge of research. Hopefully, by practicing authentic assessment well, teacher's research passion would also increase because a teacher also serves as a classroom researcher.

REFERENCES

- Akhter, S. 2017. Portofolio Assessment In Learning And Developing English Writing Skills: A Case Study Among Undergraduate Students Of Uits. *Research Journal of English Language and Literature (RJELAL)*. Vol. 5. Issue 2. (April – June). Pp 87 - 95
- Aliningsih, F. and Sofwan, A. 2015. English Teachers' Perceptions and Practices of Authentic Assessment. *Language Circle: Journal of Language and Literature X/1*, October. pp. 19-27.
- Almond, R. G. 2014. Using Automated Essay Scores as an Anchor When Equating Constructed Response Writing Tests. *International Journal of Testing*. Vol. 14, No. 1, pp. 73-91.

- Direktorat Pembinaan SMA Ditjen Pendidikan Dasar dan Menengah. 2017. *Panduan penilaian oleh Pendidik dan Satuan Pendidikan Sekolah Menengah Atas*. Direktorat Pendidikan Sekolah Menengah Atas Ditjen Pendidikan Dasar dan Menengah. Kementerian Pendidikan dan Kebudayaan. Jakarta
- Evers, A., et. al. 2017. Testing Practices and Attitudes Towards Tests and Testing: An International Survey. *International Journal of Testing*. Vol. 17, No. 2, pp. 158-190.
- Fox, James, et. al. 2017. "Keeping It Real" A Review of Benefits, Challenges and Steps Forwards Implementing Authentic Assessment. *AISHE Journal*. Vol. 9, Number 3 (Autumn 2017). pp. 3231-32313 URL: <http://ojs.aishe.org/index.php/aishe-j/article/view/280>
- Hung, L. 2016. Alternative Assessment: Can Portfolio Assessment have Positive Impact on EFL Aboriginal Students' Learning Outcome. *International Journal of Management and Applied Science*. Vol. 2, No. 9, pp. 139-144.
- Menteri Pendidikan dan Kebudayaan Republik Indonesia. 2013. Salinan Lampiran Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia Nomor 66 Tahun 2013 Tentang Standar Penilaian Pendidikan.
- Pan, Y.-C. 2016. A Review of Washback and its Pedagogical Implications. *VNU Journal of Science in Foreign Languages*, Vol. 25 No. 4, 257-263
- Reynisdottir, B. B. 2016. *The Efficacy of Authentic Assessment A Practical Approach to Second Language Testing*. Department of English School of Humanities. University of Iceland
- Sa'idah, N., Yulistianti, H. D., and Farida, Y. E. 2017. Efektivitas Penilaian Otentik Pembelajaran Bahasa Indonesia Untuk Peningkatan Kinerja Ilmiah Siswa. *Jurnal Refleksi Edukatika*, Vol. 8, No. 1. <http://Jurnal.umk.ac.id/index.php/RE>
- Siddiek, D. A. 2010. The Impact of Test Content Validity on Language Teaching and Learning. *Asian Social Sciences*. Vol. 6, No. 12, pp. 133-143.
- Sukardi. 2011. *Evaluasi Pendidikan: Prinsip dan Operasionalnya*. Jakarta: Bumi Aksara.
- Sunarti and Rahmawati, S. 2014. *Penilaian dalam Kurikulum 2013: Membantu Guru dan Calon Guru Mengetahui Langkah-Langkah Penilaian Pembelajaran*. Yogyakarta: Penerbit ANDI.