Teacher Assessment Literacy: Discrepancies in Authentic Assessment Practice in EFL Context

A.A. Istri Ngurah Marhaeni, Nyoman Dantes, and A.A.G. Yudha Paramartha
Ganesha University of Education
ngurahmarhaeni2603@gmail.com

Abstract. Teacher assessment literacy (TAL) refers to the ability of a teacher to conduct valid assessment and make use of assessment results to promote learning. As the 2015 PISA’s report indicates that Indonesian students’ achievements in three basic subjects (Reading, Science, and Math) are very low, then, a pivotal question is whether those teachers are literate enough in conducting valid assessment particularly classroom, authentic assessment. In response to such situation, this study aimed at knowing English as a Foreign Language teacher assessment literacy (EFL TAL) through investigating teachers’ authentic assessment implementation in EFL classrooms in some Senior High Schools in Bali, seeing discrepancies which occurred among the types of authentic assessment across three phases of assessment implementation (planning, executing, analyzing and reporting). Data were collected through document analysis and observation. Interviews were conducted afterwards. Findings include, 1) discrepancies found in the implementation of types of assessment ranged from moderate to high, where the highest discrepancy occurred in the practice of portfolio assessment; 2) discrepancies found across phases of assessment implementation also ranged from moderate to high, where the highest discrepancy occurred in analyzing and reporting phase. The findings imply that EFL TAL might be the source of the discrepancies.

Keywords: authentic assessment, discrepancy, teacher assessment literacy

INTRODUCTION

Among the three main tasks of a teacher (planning the lesson, executing the lesson, and assessing students’ growth), assessment is very often said to be the most challenging. According to Marhaeni et al. (2017), assessment requires teachers to analyze learning goals and their indicators and develop valid assessment tools. These professional works are relatively difficult because, up to now still many teachers think that the most important task is to teach in the class; therefore assessment is often neglected. A more fundamental reason is that because assessment needs logical and critical thinking and use some statistical computation.

Teacher assessment literacy refers to the ability of a teacher to conduct valid assessment and make use of assessment results to promote learning (Stiggins, 1994). This literacy is the minimum requirement that a teacher has to master in educational assessment in order to be effective in teaching and assessing. However, the unsatisfying results of PISA of Indonesian students show that quality of assessment practiced by teachers is still low. Some aspects of critical thinking, like predicting and analyzing which appear on the PISA tests on math, science, and reading, can not be answered correctly by Indonesian students participating in the program.

According to Stiggins (1994), teachers with assessment literacy know what they assess, why they assess, how to assess, what the possible problems with assessment are, and how to prevent them from occurring; and they also are familiar with the possible negative consequences of poor, inaccurate assessment. In relation to this, Indonesian 2013 Curriculum for primary and secondary education requires teachers to implement authentic assessment to support the goals of the curriculum which are competency based. Authentic assessment is said to be one of the best way of ensuring the assessment of competencies because authentic assessment is multi-faceted and mostly real task-based which is in accord with the nature of a competency.

In response to the importance of teacher assessment literacy to support quality education, a study on teachers’ ability in implementing authentic assessment is needed. Especially in English as a Foreign Language (EFL) learning, studies on teachers’ practices of authentic assessment (Marhaeni, Dantes, & Artini, 2015) found that although there is growing practices on authentic assessment but EFL teachers’ knowledge is just average. This condition is an indication that EFL teachers assessment literacy is still in question.
Davies (in Khadijeh and Amir, 2015) believes that language assessment literacy has three basic elements: skills, knowledge, and principles. Further they quoted Fulcher (2012) who stated that assessment literacy has three aspects. First, is knowledge, skills and abilities to design, develop, maintain or evaluate large-scale standardized and/or classroom based tests. Second, is familiarity with test processes, and awareness of principles and concepts that guide and underpin practice, including ethics and codes of practice. And third, is the ability to place knowledge, skills, processes, principles and concepts within wider historical, social, political and philosophical frameworks in order to understand why practices have arisen as they have, and to evaluate the role and impact of testing on society, institutions, and individuals. Further, Khadijeh and Amir (2015) cited a model of assessment literacy presented by Abell and Siegel (2011) and Gottheiner and Siegel (2012) which is composed of three main aspects: view of learning, assessment principles, and four areas of knowledge i.e. knowledge of assessment purposes, knowledge of assessment strategies, knowledge of assessment interpretation and action taking, and knowledge of what to assess.

Marhaeni, Dantes, & Artini (2015) also found that EFL teachers’ assessment practices are still traditional, quantitatively-oriented assessment. This can also be an indication that those teachers probably ignore the functions of assessment as a formative process of learning. In terms of language learning including in EFL in which the acquisition of language needs continuous practices, the use of formative, authentic assessment like portfolio assessment, project assessment and also self- and peer-assessment is fundamental. However, so far only a few yet EFL teachers’ practices in such issue are reported. Therefore, a study on this issue is urgent to be conducted.

Based on the background above, the purpose of the study was to investigate whether or not there was a discrepancy in teacher’s practice of authentic assessment, esp. In EFL context. The result of the study could be a parameter to discuss EFL teacher’s assessment literacy.

**METHOD**

This is a descriptive research conducted to sixteen senior high school teachers of EFL in Bali. Subjects of the study were sixteen EFL teachers of eight senior high schools randomly selected from four regencies in Bali. There four kinds of authentic assessment studied, namely self assessment, performance assessment, project assessment, and portfolio assessment. An observation sheet was developed for each of the kinds of authentic assessment, following the characteristics of each of them but it was based on a common blue print. Observation was done to teachers’ planning (assessment design in syllabus and lesson plans), Execution (assessment done during and after the teaching and learning processes), and Analysis and Reporting. Below is the blue print of Project Assessment Implementation.

| Table 1. Blueprint for Observation of Project Assessment Implementation |
|-----------------|-----------------|-----------------|
| **Planning**    | **Learning Objectives**                                      |
|                 | Are based on K-13                                          |
|                 | Considering student’s characteristics                       |
|                 | In accord with the characteristics of project assessment    |
| **Standard**    | Is based on Bloom’s Taxonomy                                |
|                 | Contains relevant domain of learning (cognitive, affective, psychomotor) |
| **Design of authentic tasks** | Related to real life problems                           |
| **Assessment criteria** | In accord with the learning objectives                  |
| **Assessment rubrics** | In accord with learning objectives                        |
|                 | Using appropriate holistic and/or analytical rubrics        |
|                 | Assessing phases of project development                     |
|                 | Assessing several relevant aspects                          |
|                 | Using descriptors for each aspect                           |
|                 | Criteria for determining student’s achievement level        |
| **Execution**   | **Formative**                                              |
|                 | Assessment in process                                      |
|                 | Assessment across phases of the project                     |
|                 | Feedback                                                    |
| **Assessment Tool** | Availability of assessment tools                        |
| **Time Management** | Time allocation for project work and assessment            |
| **Analysis and Reporting** | **Quantitative**                                         |
|                 | Scoring and weighting                                      |
|                 | Descriptive, verbal assessment to students’ performance     |
The obtained data were analyzed descriptively by comparing the observation (data obtained from observation) with 100. The criteria below was used to evaluate the level of discrepancy that occurred.

**Table 2: The Criteria of Discrepancy Level in Authentic Assessment Implementation (in per cent)**

| Range of Discrepancy | Level of Discrepancy |
|----------------------|----------------------|
| 0% < D ≤ 20%         | Very narrow          |
| 21% < D ≤ 40%        | Narrow               |
| 41% < D ≤ 60%        | Moderate             |
| 61% < D ≤ 80%        | Wide                 |
| 81% < D ≤ 100%       | Very wide            |

(Guilford in Dantes, 2014)

**FINDING AND DISCUSSION**

The data obtained from 16 EFL teachers across grades 11, 12, and 13 respectively, are shown on the table below.

**Table 3. Description of Authentic Assessment Implementation**

| NO. | Kinds of Assessment | Phases of Assessment Implementation | Average |
|-----|---------------------|------------------------------------|---------|
|     |                     | Planning                           | Executing | Analysis and Reporting |       |
|     |                     | Obs. | Disc. | Obs. | Disc. | Obs. | Disc. | Obs. | Disc. |
| 1.  | Self Assessment     | 42.67 | 57.33 | 40.67 | 59.33 | 40.67 | 59.33 | 41.34 | 58.66  |
| 2.  | Performance Assessment | 42   | 58   | 39.67 | 60.33 | 40.33 | 59.67 | 40.67 | 59.33  |
| 3.  | Project Assessment  | 46   | 54   | 33.67 | 66.33 | 37.67 | 62.33 | 39.11 | 60.89  |
| 4.  | Portfolio Assessment | 39   | 61   | 32   | 68   | 34.67 | 65.33 | 35.22 | 64.78  |
| 5.  | Average             | 42.42 | 57.58 | 36.50 | 63.50 | 38.34 | 61.66 | 39.01 | 60.91  |

The table above clearly shows the data distribution of authentic assessment implementation by EFL teachers, in which there were four kinds of authentic assessment studied (Self Assessment, Performance Assessment, Project Assessment, and Portfolio Assessment) across the three phases of assessment implementation (Planning, Executing, Analyzing and Reporting).

**Question #1:** What were the discrepancies among the four kinds of authentic assessment?

The table above shows that, the average score of observation (Obs.) of all the four kinds of authentic assessment was 39.01 and the discrepancy was 60.91 which is categorized as wide. Among the four kinds of authentic assessment, the highest observation score 41.34 was found on Self Assessment, and the discrepancy was 58.66 (moderate). On the other hand, the lowest observation score 35.22 was found on Portfolio Assessment, and the discrepancy was 64.78 (wide).

**Question #2:** What were the discrepancies of authentic assessment among the phases of assessment implementation?

The table above clearly shows that the average observation score of Phases of Authentic Assessment Implementation was 39.01 and the discrepancy was 60.91 (wide). Furthermore, the observation score of Planning Phase was 42.42 and the discrepancy score was 67.58 (wide); the observation score of Execution Phase was 36.50 and the discrepancy was 63.50 (wide); and the observation score of Analysis and reporting Phase was 38.39 and the discrepancy was 61.66 (wide), respectively.

Based on the findings above, it can be clearly seen that discrepancies really occur in the implementation or practice of authentic assessment in EFL classrooms, especially in senior high schools in Bali. The discrepancy found was between moderate to wide. This means that the implementation of authentic assessment in EFL classrooms can be said not satisfying. The discrepancies also become a focus of some recent research in Indonesian education context. Abidin (2012) found discrepancy in teachers’ planning of lessons; he revealed that 68 percent teachers did not include authentic assessment in their lesson plans. In accordance with this, Zaim (2013) found that teachers spent a lot of time in planning and analyzing data, which was due to their lack of knowledge about authentic assessment. It is true what Nitko (2012) said that over fifty percent time spent by teachers to do assessment related activities, but if time spent is too long because of lack of knowledge, then, other aspects of assessment practice and also instruction in general, would face a problem in time. More specifically, Putri and Hariyati (2015) reported that teachers did not know how to design assessment rubrics. While a rubric is an essential tool in assessing students performance in authentic assessment, this lack of assessment literacy aspect on teachers inevitably
challenges quality of classroom assessment practices.

Therefore, those discrepancies as reported in this study and also other previous studies are potential to create challenges to the mandate of our 2013 curriculum, that is, through the Ministry of National Education Regulation No. 023 year 2016 about Assessment Standard, it is mandated that authentic assessment must be one of the major assessment to be implemented in instructional processes in classes. Johnson and Johnson (2002) emphasize that assessment is a process of collecting information about students’ growth and progress in learning. If assessment is not properly implemented, then the decision made based on such assessment is not valid and obsolete.

Results of this research has an implication toward the condition of teachers’ assessment literacy, in the context of EFL teaching in Bali province. Although seen only from the aspect of authentic assessment, these findings show that teachers need to improve their competency in handling meaningful assessment of students. More importantly is because, in classroom assessment, authentic assessment is the framework of teachers’ assessment practice. A good implementation of authentic assessment is needed to assure that target competencies are reached.

CONCLUSION

Based on the findings and discussion above, it is concluded that EFL teachers’ implementation of authentic assessment is not yet satisfying, seen from both the kinds of authentic assessment and across phases of assessment implementation. This result implies that EFL teachers assessment literacy needs to be improved especially in terms of knowledge and also skills in implementing authentic assessment.

*****

REFERENCES

Abidin, Y. (2012). Model penilaian otentik dalam pembelajaran membaca pemahaman berorientasi pendidikan karakter. Jurnal Pendidikan Karakter, Vol. II, Nomor 2, Hlm. 1-15

Dantes, N. (2014). Metode Penelitian. Yogyakarta: Andi Publisher.

Fazilla, S. (2013). Penerapan asesmen portofolio dalam penilaian hasil belajar sains SD. Jurnal UPI. pp. 139-154.

Fencl, M. J. (2014). Fun and Creative Unit Assessment Ideas for All Students in Physical Education. Journal of Physical Education, Recreation & Dance, 8(1), pp. 16-21.

Johnson, D. W. & Johnson R. T. (2002). Meaningful Assessment: A Manageable and Cooperative Process. Boston: Allyn and Bacon.

Kemendikbud. (2013c). Peraturan menteri pendidikan dan kebudayaan republik indonesia nomor 81a tahun 2013 lampiran 4 about pedoman umum pembelajaran. Jakarta: Kementerian Pendidikan dan Kebudayaan RI.

Kemendikbud. (2015). Peraturan menteri pendidikan dan kebudayaan nomor 53 tahun 2015 about penilaian hasil belajar oleh pendidik dan satuan pendidikan pada pendidikan dasar dan menengah. Jakarta: Kementerian Pendidikan dan Kebudayaan RI.

Khadijeh, B. & Amir, R. (2015). “Importance of Teachers Assessment Literacy”. International Journal of English Language Education, Vol. 3; No. 1 ISSN 2325-0887

Maharani, D., N. (2013). The Effect of Self-Assessment on Students Writing Skills across the Self-Efficacy. Jurnal Pendidikan Bahasa Inggris(1). pp. 1-12.

Marhaeni, A.A.I.N., Dantes, N., & Artini, L.P. (2015). Asesmen Autentik dan Pendidikan Bermakna: Implementasi Kurikulum 2013. Jurnal Pendidikan Indonesia. 4(1), April 2015.

Marhaeni, A. I. N. et al. (2017). Asesmen Autentik dalam Pembelajaran Bahasa Inggris. Jakarta: Rajagrafindo.

Nuraffifah. (2014). Penerapan Asesmen Kinerja untuk Menumbuhkan Keterampilan Menganalisis Siswa pada Pembelajaran Sejarah di Kelas XI IPS 2 SMAN 1 Kota Bandung. Jurnal UPI. pp. 1-11.

O’Malley, J.M. & Valdez Pierce, L. (1996). Authentic Assessment for English Language Learners.New York: Addison-Wesley Publishing Company.

Rahmat (2003). Undang-undang Republik Indonesia Nomor 20 tahun 2003: Sistem Pendidikan Nasional. Sekretaris Negara Republik Indonesia.

Putri, W.N.A. & Hariyati. (2015). Implementasi Penilaian Portofolio Guru Mata Pelajaran IPA di SMP Negeri 2 Banyudono. Unis E-print Journal. pp. 1-12.

Stiggins, R. J. (1994). Student Centered Classroom Assessment. New York: Maxwell Macmillan International

Theresia, T., P. (2015). The Implementation of Self-Assessment in Writing Class: a Case Study at STBA LIA Jakarta. TEFLIN Journal 26(1), pp. 97-116. Doi: dx.doi.org/10.15639/teflinjournal.v26i1

Zaim, M. (2013). Asesmen Otentik: Implementasi Dan Permasalahannya Dalam Pembelajaran Bahasa Inggris Di Sekolah Menengah. Proceeding International Seminars on Language and Arts FBS Universitas Negeri Padang, pp. 39-64.