English Language Teachers' Uses of Classroom Assessment

Hussain Alkharusi
Sultan Qaboos University, Oman, *hussein5@squ.edu.om*

This study aimed at identifying the purposes of using assessment by English language teachers in Oman. Participants were 369 English language teachers randomly selected from all governorates in Oman. The teachers were categorized into two groups based on the assessment system. The first group consisted of teachers teaching grades 7 to 9 because the assessment system in these grades places more emphasis on formative assessment. The second group consisted of teachers teaching grades 10 and 11 because the assessment system in these grades places more emphasis on summative assessment. A questionnaire was developed and administered to all teachers regarding the purposes of using assessment. A factor analysis of the teachers' responses revealed two dimensions of the assessment purposes: Instructional- versus learning-oriented purposes. Instructional-oriented purposes were related to teachers' use of assessment for evaluating their teaching methods and managing student behavior. Learning-oriented purposes were related to teachers' use of assessment for motivating students to learn and assigning grades for students. There were significant differences on using assessment for instructional-vs-learning-oriented purposes favoring learning-oriented purposes. There were no significant grade level differences on using assessment for instructional-vs-learning-oriented purposes. These results lead to a conclusion that the assessment system might not have an impact on the purposes for which assessment is used by the teachers.

Keywords: classroom assessment, assessment purposes, instruction, learning, English language teachers

INTRODUCTION

Classroom assessment aims at improving student learning and motivation to learn (Gronlund, 2006). It has become a tool for improving classroom teaching and learning (Shavelson, Young, Ayala, Brandon, Furtak, Ruiz-Primo, Tomita, & Yin, 2008). In this regard, Gronlund (2006) suggests that a sound classroom assessment requires a clear conception of all intended learning outcomes of the instruction and a variety of assessment procedures that are relevant to the instruction, adequately sample student performance, and fair to everyone. In addition, a sound assessment requires the specifications of criteria for judging successful performance and timely and detailed feedback to students emphasizing strengths of their performance and weaknesses to be corrected (Gronlund, 2006).

Teachers are required to develop classroom assessment that aligns with practices recommended by experts of educational assessment. For example, assessment experts have recommended that students should clearly be informed about the grading procedure in advance and involved in the assessment process; student personal characteristics such as ability, effort, motivation, interest, and neatness of work should not be incorporated into grading due to the lack of objective measurement; a final grade for borderline cases should be determined using additional academic achievement data rather than non-achievement data; and students should be given continuous and informative assessment feedback rather than judgmental feedback about their academic performance (Brookhart, 1994; Gronlund, 2006; Stiggins & Chapuis, 2005).

Moreover, teachers are encouraged to use more than one assessment method in order to have enough, accurate evidence of student learning (Nitko, 2001). It is emphasized that the assessment should match the learning targets and provides meaningful feedback to students (Nitko, 2001). In addition, the American Federation of Teachers (AFT), the National Council on Measurement in Education (NCME), and the National Education Association (NEA) (1990) have jointly defined seven Standards for Teacher Competence in Educational Assessment of Students. The standards emphasized that teachers should competently be able to choose and develop assessment methods appropriate for instructional decisions; administer, score, and interpret results of externally produced and teacher-made assessment; use assessment results when making educational decisions; develop valid assessment-based grading procedures; communicate assessment results; and recognize unethical, illegal, and inappropriate methods and uses of assessment (AFT, NCME, & NEA, 1990).

English language teachers use classroom assessment for many purposes. These include evaluating teaching methods, motivating students to learn, comparing students' performances, reporting grades to students, parents and other stakeholders (Jabbarifar, 2009). Using a mixed-methods research design with 48 Filipino and Indonesian English language teachers, Saefurrohman and Balinas (2016) found that teachers used assessment for learning as the primarily purpose of classroom assessment. Further, using a participatory action research with six language teachers in Uganda, Akello and Timmerman (2018) found that using assessment of learning is more effective in small classes than in large classes. Also, large classes and teachers' workload might be threats to the appropriate use of the classroom assessment. In a quantitative descriptive research study of 101 secondary English language teachers in Iraq, Muhammad and Bardakçi (2019) reported that teachers are not well prepared for using classroom assessment either for formative or summative purposes. The use of classroom assessment might be affected by a number of factors including the assessment policy and grade level (Liu, 2008). Recently, Narathakoon, Sapsirin, Subphadoongchone (2020) added that the discrepancies between teachers' classroom assessment beliefs and practices might be affected by other factors such as the educational policy, limitation of time, school workload, and knowledge of assessment.

When considering the assessment context in Oman, the grading weight given to formative versus summative assessment differs by grade level. The formative assessment in grades 7 to 9 is given 60% out of the total student mark in English language subject whereas the summative assessment is given 40%. In contrast, the formative assessment in grades 10 and 11 is given 40% out of the total student mark in English language subject whereas the summative assessment is given 60%. Therefore, the research problem being addressed in this study was that whether differences in the assessment regulations of the English language subject across grade levels make differences in teachers' use of classroom assessment.

Research Questions

Specifically, the study aimed to answer the following research questions:

1.   What are the underlying purposes of using classroom assessment by English language teachers in grades 7 to 11 in the Sultanate of Oman?
2.   Are there statistically significant differences on the underlying purposes of using classroom assessment by English language teachers in grades 7 to 11 in the Sultanate of Oman?
3.   Are there statistically significant grade level differences on the underlying purposes of using classroom assessment by English language teachers in the Sultanate of Oman?

METHOD

Research Design

This study employed a descriptive survey research design that is cross-sectional in nature to describe the purposes of using classroom assessment by English language teachers in grades 7 to 11 in Oman.

The following sections provide more details about the participants, data collection process, instrumentation, and analyses.

Participants

The participants in this study were 369 English language teachers. They were randomly selected from all educational regions in Oman. There were 148 males and 221 females. They were teaching grades 7 to 11. The distribution of the teachers across grade levels was as follows: 58 (15.7%) grade 7, 58 (15.7%) grade 8, 63 (17.1%) grade 9, 81 (22%) grade 10, 109 (29.5%) grade 11. Their teaching experience ranged from 2 to 20 years with an average of 10 years and a standard deviation of 7. The majority of them (91.6%) had an undergraduate degree in Education whereas the rest had a postgraduate degree in Education.

Research Procedures

After determining the research questions to be investigated, the appropriate research design of the study was determined based on the purposes of the study. Then, a literature search was conducted to find relevant studies. Also, the literature was consulted to develop the appropriate instrument to collect the needed data for answering the research questions. After developing the instrument, it was subjected to a content validation process. Information about the sample was obtained from the Ministry of Education. Then, the data collection process began as follows.

Permission was requested from Ministry of Education, schools' administration, and teachers to collect data from English language teachers teaching grades 7 to 11. The participants were informed that a study is being conducted to investigate their uses of classroom assessment. The teachers were also informed that they were not obligated to participate in the study, and that if they wished, their responses would remain anonymous and confidential. Those who wished to participate in the study were provided a cover letter and a questionnaire, which is described in the next section. Finally, the collected data were prepared for the statistical analysis.

Instrument

Informed by the literature (Gallagher, 1998; Gronlund, 2006; Nitko, 2001; Poham, 2000), a 10-item questionnaire was developed and used in this study. The items were concerned about the uses of classroom assessment: evaluating effectiveness of teaching methods, diagnosing strengths and weakness of student learning, grouping students for instruction, determining student readiness for new learning, assigning grades to students, controlling student behaviours, motivating students to learn, evaluating student achievement, comparing student achievement across classes, and transferring students to a higher grade level. The participants were asked to indicate the extent to which they use classroom assessment for the purposes described in the item on a 5-point Likert scale ranging from 1 (never) to 5 (always). To establish content validity, the items were given to a group of faculty members in the areas of educational measurement and psychology from Sultan Qaboos University. They were asked to judge the clarity of wording and the appropriateness of each item and its relevance to the construct being measured. Their feedback was used for further refinement of the items.

Data Analysis

An exploratory factor analysis was performed to identify the underlying purposes of using classroom assessment by English language teachers in grades 7 to 11 in the Sultanate of Oman. The factor analysis was used because it is the appropriate statistical technique to reflect the underlying process and pattern by using observed variables (Tabachnick & Fidell, 2001). Internal consistency reliability for each factor was established through Cronbach's alpha. Then, within-and-between-subjects ANOVA was performed to examine differences on the underlying purposes of using classroom assessment by English language teachers in grades 7 to 11 in the Sultanate of Oman and to investigate grade level differences with respect to the purposes of using classroom assessment. The within-

subjects factor was the type of purposes of using classroom assessment. The between-subjects factor was the grade level. The teachers were categorized into two groups based on the assessment system. The first group consisted of teachers teaching grades 7 to 9 because the assessment system in these grades places more emphasis on formative assessment compared to summative assessment. The second group consisted of teachers teaching grades 10 and 11 because the assessment system in these grades places more emphasis on summative assessment compared to formative assessment.

FINDINGS

The participants' responses were submitted to the principal components analysis to identify their underlying dimensions. No particular number of dimensions was hypothesized and the criterion was set to eigenvalues greater than one (Tabachnick & Fidell, 2001). The Bartlett's test of sphericity was statistically significant, $\chi^2(45) = 979.28$, $p < .001$. The Kaiser's measure of sampling adequacy (MSA) fell within acceptable range (values of .60 and above) with a value of .82. Each item also exceeded the threshold value (.60) of MSA. Finally, the partial correlations were small as indicated by the anti-image correlation matrix. These measures all led to the conclusion that the set of 10 items of the classroom assessment purposes was appropriate for principal components analysis. The initial unrotated principal components analyses resulted in a factor model of two dimensions as suggested by the scree plot and eigenvalues exceeding unity. However, based on its pattern of factor loadings, this unrotated factor model was theoretically less meaningful and difficult to interpret. Therefore, the analysis proceeded to rotate the factor matrix orthogonally with Varimax rotation to achieve a simple and theoretically more meaningful solution.

As suggested by the eigenvalue rule and scree plot, the analysis yielded a two-factor structure accounting for 49.91% of the total variance. The first factor accounted for 37.85% of the total variance (eigenvalue = 3.785) and consisted of five items with factor loadings ranging between .56 and .78. Based on the content of its items, this factor was named "instructional-oriented purposes" because the items were related to teachers' use of assessment for evaluating teaching methods and managing student behavior. The second factor accounted for 12.06% of the total variance (eigenvalue = 1.206) and consisted of five items with factor loadings ranging between .45 and .80. Based on the content of its items, this factor was named "learning-oriented purposes" because the items were related to teachers' use of assessment for motivating students to learn and assigning grades for students. Table 1 presents the items, the factor loadings, the percent of the variance explained, and Cronbach alpha reliability for each factor.

Table 1
Factor structure of the classroom assessment purposes

Items	Factor loadings	
	F1	F2
1. Controlling student behaviors	.78	
2. Determining student readiness for new learning	.71	
3. Grouping students for instruction	.62	
4. Evaluating effectiveness of teaching methods	.62	
5. Comparing student achievement across classes	.56	
6. Evaluating student achievement		.80
7. Assigning grades to students		.79
8. Motivating students to learn		.57
9. Diagnosing strengths and weakness of student learning		.47
10. Transferring students to a higher grade level		.45
% of variance explained	37.85%	12.05%
Reliability	.72	.82

Note. F1 = instructional-oriented purposes. F2 = learning-oriented purposes.

Based on the principal components analyses, two scales were formed for the classroom assessment purposes. Each scale score was created by averaging the individual scores of all items loading on a factor. A high scale score reflected a high use of assessment purposes described by the scale.

Within-and-Between-Subjects ANOVA was conducted to examine differences on the purposes of using classroom assessment by English language teachers in grades 7 to 11 in the Sultanate of Oman and to investigate grade level differences with respect to the purposes of using classroom assessment. Table 2 shows means and standard deviations of using classroom assessment with respect to grade level and type of purpose.

Table 2
Means and standard deviations of using classroom assessment with respect to grade level and type of purpose

Variables	n	M	SE	*95% CI*
Instructional-oriented purposes				
Grades 7 – 9	179	3.92	.05	[3.82, 4.02]
Grades 10 – 11	190	3.79	.05	[3.70, 3.89]
Learning-oriented purposes				
Grades 7 – 9	179	4.41	.04	[4.34, 4.49]
Grades 10 – 11	190	4.36	.04	[4.29, 4.44]

Results of the within-subjects analysis revealed statistically significant differences among teachers on using assessment for instructional-vs-learning-oriented purposes; $F(1, 367) = 317.945$, $p = .000$; partial $\eta^2 = .46$. Teachers tended on average to use classroom assessment for learning purposes ($M = 4.39$; $SE = .03$; 95% CI [4.34, 4.44]) more than for instructional purposes ($M = 3.85$; $SE = .04$; 95% CI [3.79, 3.93]). Results of the between-subjects analysis revealed no statistical significant grade level differences on using assessment for instructional-vs-learning-oriented purposes; $F(1, 367) = 2.522$, $p = .113$; partial $\eta^2 = .007$.

DISCUSSION, CONCLUSION AND SUGGESTIONS

The present study investigated the purposes of using classroom assessment by English language teachers in grades 7 to 11 in the Sultanate of Oman. A factor analysis of the teachers' responses revealed two dimensions of the assessment purposes: Instructional- versus learning-oriented purposes. Instructional-oriented purposes were related to teachers' use of assessment for evaluating their teaching methods and managing student behaviour. Learning-oriented purposes were related to teachers' use of assessment for motivating students to learn and assigning grades for students. There were significant differences on using assessment for instructional-vs-learning-oriented purposes favouring learning-oriented purposes. There were no significant grade level differences on using assessment for instructional-vs-learning-oriented purposes. These results imply that the primary purpose of using classroom assessment by English language teachers in grades 7 to 11 in the Sultanate of Oman is to improve students' learning rather than to improve teachers' teaching. Specifically, assessment in this context is mainly used for describing, judging, and reporting student performance. It is less frequently used for making pedagogical decisions. Although the need to provide feedback concerning student's attainment of the instructional objectives is important, teachers should make reflections on their instructional practices based on the assessment results (Stiggins, 1992). The current findings add support to the previous studies in terms of conducting classroom assessment by teachers as assessment of learning more than as assessment for learning (Akello & Timmerman, 2018; Narathakoon et al., 2020; Saefurrohman & Balinas, 2016). Also, the findings stressed the importance of offering professional development programs for teachers on how to use classroom assessment for various purposes.

This study is a small step towards a complete understanding of how do English language teachers in the Sultanate of Oman deal with classroom assessment. Qualitative research approaches might shed

more light on teachers' conceptions of assessment and reflections on their assessment practices. Also, future studies might consider conducting a confirmatory factor analysis to test the validity of the factorial structure underlying the classroom assessment purposes identified in the present study. The study attempted to examine whether the assessment regulation of the Ministry of Education at Sultanate of Oman makes a difference in teachers' uses of assessment. The results lead to a conclusion that despite the differences in the assessment regulation, assessment in English language classes is used primarily for making decisions about student performance more than for evaluating the soundness of instructional practices.

REFERENCES

Akello, L. D., & Timmerman, M. C. G. (2018). Formative assessment: The role of participatory action research in blending policy and practice in Uganda. *Educational Action Research, 26*, 736 – 754. https://doi.org/10.1080/09650792.2017.1405831

American Foundation of Teachers, National Council on Measurement in Education, & National Education Association. (1990). Standards for teacher competence in educational assessment of students. *Educational Measurement: Issues and Practice, 2*, 30 – 32. https://doi.org/10.1111/j.1745-3992.1990.tb00391.x

Brookhart, S. M. (1994). Teachers' grading: Practice and theory. *Applied Measurement in Education, 7*, 279 – 301. https://doi.org/10.1207/s15324818ame0704_2

Gallagher, D., Jo. (1998). *Classroom assessment for teachers*. New Jersey, NJ: Prentice-Hall.

Gronlund, N. E. (2006). *Assessment of student achievement* (8th ed.). Boston: Pearson.

Jabbarifar, T. (2009). *The importance of classroom assessment and evaluation in educational system*. Proceedings of the 2nd International Conference of Teaching and Learning, INTI University College, Malaysia.

Muhammad, F. H. N., & Bardakçi, M. (2019). Iraqi EFL teachers' assessment literacy: Perceptions and practices. *Arab World English Journal, 10*, 420 – 430. https://doi.org/10.24093/awej/vol10no2.33

Narathakoon, A., Sapsirin, S., & Subphadoongchone, P. (2020). Beliefs and classroom assessment practices of English teachers in primary schools in Thailand. *International Journal of Instruction, 13*, 137- 156. https://doi.org/10.29333/iji.2020.13310a

Nitko, A. J. (2001). *Educational assessment of students* (3rd ed.). Upper Saddle River, NJ: Prentice-Hall.

Saefurrohman, & Balinas, E. S. (2016). English teachers classroom assessment practices. *International Journal of Evaluation and Research in Education, 5*, 82 – 92. https://doi.org/10.11591/ijere.v5i1.4526

Shavelson, R. J., Young, D. B., Ayala, C. C., Brandon, P. R., Furtak, E. M., Ruiz-Primo, M. A., Tomita, M. K., & Yin, Y. (2008). On the impact of curriculum-embedded formative assessment on learning: A collaboration between curriculum and assessment developers. *Applied Measurement in Education, 21*, 295 – 314. https://doi.org/10.1080/08957340802347647

Stiggins, R. J. (1992). High quality classroom assessment: What does it really mean? *Educational Measurement: Issues and Practice, 11*, 35 – 38. https://doi.org/10.1111/j.1745-3992.1992.tb00241.x

Stiggins, R., & Chappuis, J. (2005). Using student-involved classroom assessment to close achievement gaps. *Theory Into Practice, 44*, 11 – 18. https://doi.org/10.1207/s15430421tip4401_3

Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.). Needham Heights, MA: Allyn & Bacon.

O'Neill, G., & Murphy, F. (2010). Guide to taxonomies of learning. *UCD Teaching and Learning/Resources*, Retrieved 01 November, 2010 from http://www.ucd.ie/t4cms/ucdtla0034.pdf