

RESEARCH

Open Access



# Assessment approaches of English language teachers in the Saudi higher education context

Samar Yakoob Almosa\* and Sahar Matar Alzahrani

\* Correspondence: [symossa@uqu.edu.sa](mailto:symossa@uqu.edu.sa)  
English Language Centre, Umm Al-Qura University, Makkah, Saudi Arabia

## Abstract

Assessment approaches including assessment purposes, assessment processes, fairness, and measurement theory, and English teachers' professional development needs remain underexplored in the Middle East and North Africa regions. This study provided empirical evidence on English language teachers approaches in the Saudi higher education context. A survey was used to examine the teachers' current approaches to classroom assessment. A total of 287 subjects (191 women and 94 men) participated in the survey. The results revealed that both the male and female participants valued and endorsed assessments alike. However, female participants were found to value assessment purposes more than their male counterparts. Fairness in assessment approaches was the least valued item in teachers' identified assessment approaches. Experienced teachers who identified themselves as competent in their role valued assessment fairness and measurement theory more than novice teachers. The present work broadens our knowledge on teachers' assessment approaches in relation to gender, career stage, and academic position, which support interested researchers and policy-makers in decision-making regarding designing professional development programs.

**Keywords:** English teachers, Higher education, Assessment approaches

## Introduction

Assessment is believed to be a cornerstone in the educational system that plays an integral role in impacting various processes like teaching, learning, and decision-making (Coombs et al. 2018). Specifically, assessment to gauge accountability is a major concern in general and in higher education alike (Deluca et al. 2016a). To meet the accountability demands, work has been conducted to ensure that institutions meet the requisite standards and respond to demands that stress the role of assessment in informing policy and practices (Deluca et al. 2018).

Along with acknowledging the positive or negative impact of teachers' approaches to assessment on students' classroom learning (Harlen, 2006; Hattie, 2008; Cauley and McMillan, 2010), we must also acknowledge that teachers' approaches reflect their

beliefs and views regarding teaching and learning (Xu and Brown, 2016; Looney et al. 2018; Herppich et al. 2018).

Assessment literacy and identifying the assessment skills and knowledge of teachers and stakeholder groups has become a matter of concern (Popham, 2013; Willis et al. 2013; Xu and Brown, 2016; DeLuca et al. 2018). While assessment literacy is an area that has received growing attention, it requires further theoretical investigation (Deluca et al. 2018) in school and higher education contexts. Current theories regarding assessment literacy go beyond focusing on teachers' sets of skills and knowledge and consider various sources of knowledge that shape teachers' assessment approaches, such as context and experience (Herppich et al. 2018). Various factors could influence teachers' assessment approaches and preferred methods of assessment, including assessment education needs and preferences (Coombs et al. 2018). Therefore, there is a need to closely explore teachers' assessment approaches, training needs, and training method preferences in different educational contexts. This work will pave way for a better understanding of assessment approaches. This study focused on teachers' assessment approaches to gain an insight into their thinking and approaches and how they align with contemporary assessment standards. The information gained from the teachers about their professional development in assessment education needs (assessment literacy) and their training preferences (methods of training), assessment approaches, assessment PD needs, and preferences will help policy-makers in making decisions related to assessment practices. The data collected also provided additional insights into the links between English teachers' demographic data and assessment approaches and the impact of these variables on training needs and preferences. The research implications provide information on assessment literacy development in the EFL context.

This study intends to fill the literature gap between the assessment approaches and PD needs of English language teachers in Saudi Arabia. This study specifically examines the following research questions:

1. What are the assessment approaches used by English language teachers in Saudi universities?
2. What are English teaching staff assessment training needs and preferred methods of training?
3. What is the impact of the demographic characteristics of English teaching staff on their assessment approaches?
4. What is the impact of the demographic characteristics of English teaching staff on their professional development in assessment needs?

## **Literature review**

### **Teachers' assessment approaches**

Research on measuring teachers' approaches to assessment provides useful data to support teachers' assessment literacy initiatives. Deluca et al. (2016a) stressed the importance of measuring teachers' assessment literacy in light of contemporary assessment standards that focus on demand to inform policy and practice. Several measurements for assessment literacy have been developed based on 1990 standards. Deluca et al. (2016a) analyzed assessment literacy standards developed in 1990 from Australia,

Canada, New Zealand, UK, USA, and Mainland Europe, including 14 assessment standards and eight measures, and found that the measures were based on an early understanding of assessment literacy concepts. Gotch and French, 2014 noted that some assessment measurements are problematic and no longer match current modern assessment demands. There is a lack of reliable data on teachers' assessment approaches in the Western context (DeLuca et al. 2016a; Gotch and French, 2014) and data and empirical research evidence in the non-Western educational context, such as in Middle East and North Africa (MENA) (Almoosa, 2018). Teachers' approaches to assessment are influenced by their conceptualization and practical knowledge, constructed from their educational context (DeLuca et al. 2018).

### Exploring assessment literacy

The majority of previous assessment literacy surveys were based on 1990 standards for teacher competency among educational assessment students. Gotch and French (2014) conducted a systematic review of 36 assessment literacy measures and found that the measures did not support psychometric properties and lacked representativeness and relevance of content in light of transformation in the assessment landscape. Building on these findings, demand for assessment literacy measurements that meet current assessment requirements increased. In the same vein, Brookhart (2011) argued that 1990 assessment standards no longer reflect the needed assessment knowledge teachers are expected to have or the assessment approaches to be acquired in modern classrooms. DeLuca et al. (2016b), in response to these demands, developed the Approaches to Classroom Assessment Inventory (ACAI), which reflects the latest version of 1990 classroom assessment standards. The ACAI is a two-part survey addressing teachers' approaches to classroom assessment, which includes a demographic section and scenario-based questions followed by a series of common assessment responsibilities aligned with contemporary assessment standards. In addition, the ACAI contains questions related to assessment training and preferred methods for professional assessment education. ACAI questions were developed based on a four-dimensional framework for assessment literacy predicated on analysis of 15 contemporary assessment standards, from 1990 to present, from five geographic regions: USA, Canada, UK, Europe, Australia, and New Zealand (DeLuca et al. 2016a). The four assessment dimensions include purposes, processes, fairness, and theory. Three priority areas (i.e., assessment approaches) for each dimension were selected as shown in Table 1.

Previous studies explored teachers' conceptualization of assessment purposes (e.g., Brown, 2004; Barnes et al. 2017), development of assessment literacy (Brown, 2004; DeLuca et al. 2016a; Coombs et al. 2018; Herppich et al. 2018), and specific classroom assessment approaches (Cizek et al. 1995; Cauley and McMillan, 2010). Stemming from the assumption that teachers' assessment actions have a significant influence on students' learning experience and achievement (Black and Wiliam, 1998; Hattie, 2008; DeLuca et al. 2018), there is a need to understand differences and similarity in teachers' approaches to assessment across various learning and teaching contexts (Willis et al., 2013).

**Table 1** ACAL assessment dimensions and sets of priorities

Theme	Priority	Description of priority
Assessment purposes	Assessment of learning	Teachers use of evidence to summate student learning and assign a grade in relation to the student's achievement of learning objectives.
	Assessment for learning	Teachers' and students' use of evidence to provide feedback on progress toward learning objectives and inform next steps for learning and instruction. Involves both teacher-directed and student-centered approaches to formative assessment.
	Assessment as learning	Focuses on how the student is learning by providing feedback or experiences that foster students' metacognitive abilities and learning skills, such as self-assessment, goal-setting, and learning plans. Involves teachers but is primarily student-centred.
Assessment processes	Design	Focuses on the development of reliable assessments and items that measure student learning in relation to learning objectives.
	Scoring	Focuses on the adjustment and use of scoring protocols and grading schemes to respond to assessment scenarios.
	Communication	Focuses on the interpretation of assessment results and feedback through communication to students and parents.
Fairness	Standardized	Maintains equal assessment protocols for all students regardless of ability or exceptionality.
	Equitable	Differentiates assessment protocols for formally identified students, such as special education or English language learners.
	Differentiated	Individualizes learning opportunities and assessments that address each student's unique learning needs and goals.
Assessment theory	Reliability	Works to ensure consistency in results within assessments, across time periods, and between teachers.
	Validity	Works to ensure that the assessment or evaluation measures what it claims to measure (i.e., learning objectives) and promote valid interpretations of results.
	Mixed	Works to ensure consistency in measuring what an assessment or evaluation intends to measure, and degree to which an assessment or evaluation measures what it claims to measure.

### Assessment literacy in Saudi EFL context

Despite the importance of EFL teachers' assessment literacy in higher education, very few studies have explored this topic within the Saudi Arabian context (Almoossa, 2018; Ezza, 2017; Hakim, 2015; Rauf and McCallum, 2020; Umer et al. 2018). Almoossa (2018) explored the classroom-based assessment practices of six language teachers and found a disparity between the teachers' conceptualization and their actual daily practices. She also reported that English teachers in Saudi region lack adequate pre-service and in-service training related to classroom-based assessment practices. Along similar lines, Hakim (2015) explored English teachers' levels of assessment literacy in the language centre of a Saudi university and reported that the teachers exhibited inadequate classroom-based assessment practices despite their knowledge about assessment principles and techniques. Almoosa (2021) noted that English language institutes and centres in Saudi universities followed a unified system for assessment that focused heavily on testing (examinations). The unified system limited the teachers' options and their potential for learning and developing their assessment literacy given the limited roles they played in assessment (Almoossa, 2018, Almansory, 2016). Rauf and McCallum (2020) examined writing assessment tasks performed by English teachers in relation to assessment principle and learning outcomes. They concluded that there exists a disparity between assessment principles and the practices that the participants used. They also reported that the tasks were focused on basic skill levels; similar to the results obtained

by Umer, Farooq, and Gulzar, 2018 who reported that English teachers' assessment practices were not in tandem with the learning outcomes. They noted that much emphasis was put on memorization and recalling information rather than on higher order learning outcomes.

### **Methodology**

For the purpose of this study, the ACAI survey was adopted with some modifications in the present study, as it has simple, clear, and direct statements, which are easy to understand for teachers with no or limited assessment experience. This study aimed to ensure the survey reflected contemporary assessment standards while being accessible to teachers with no jargon. It targeted teachers who may have no previous training in teaching English or came from other fields. In addition, it matched the core objective of this investigation, which focused on English language teachers' approaches to assessment purposes, process, fairness, and measurement theory.

For the current study, a few modifications were made to the original survey to suit the study population and the research study purpose. Few statements were modified to fit the university context practices and terms in the university context were also used. The section on the scenario part was omitted based upon the feedback received during the pilot study phase.

Part one of the survey consisted of demographic information related to the participants gender, age, education, job, years of experience, experience in the current role (novice, competent, expert), education, and role in assessment at their institution. Part two of the survey included statements on the various assessment approach statements, followed by part three that included information about PD preferences and preferred methods of training.

### **Demographic summary**

The study included 287 participants (191 men and 94 women). The participants were teachers in English language centres in Saudi universities, including teachers from Saudi Arabia and other countries. The survey was distributed online and the teachers were invited through official contact with university administration, personal contact by email, and twitter private invitation. The majority of the participants had a minimum of 6–9 years of teaching experience (Table 2).

### **Data analysis**

Data from the second part of the ACAI, which asked participants to identify their level of agreement with statements related to assessment tasks and responsibilities, were analyzed. Quantitative analyses included descriptive statistics, exploratory factor analysis, one-way analysis of variance, independent samples *t* test, and chi-square tests. Exploratory factor analysis was used to uncover the underlying factor structure of items. One-way analysis of variance and independent samples *t* test were used to identify statistical differences in factor scores between demographic groups. Chi-square tests were used to identify statistical differences in preferred methods of assessment education and demographic groupings. All analyses were conducted using SPSS statistical software.

**Table 2** Participant demographics

<b>Demographic variable</b>		<b>Number of participant teachers</b>
Gender	Female	191
	Male	94
Age	22–29	52
	30–39	142
	40–49	69
	50+	22
Job title	Assistant professor	48
	Associate professor	8
	Lecturer	94
	Instructor	99
	Teaching assistant	18
	Other	18
Education	PhD	70
	MA	178
	BA	40
	Other	13
Degrees	Applied linguistics	68
	TESOL	55
	Linguistics	30
	Education	32
	Literature	35
Career stage	0–2	56
	3–5	59
	6–8	41
	9+	128
Experience in current role	Novice	35
	Competent	139
	Expert	111
Assessment education	A course	162
	No course	113
Assessment decision-maker	Yes	59
	No	162
Roles in the profession	Language teacher	271
	Professional examiner and/or rater	35
	Member of test development unit	45
	Researcher of language assessment	27
	Policy-maker for language assessment	15
	Test score user	0

## Results

### RQ1: What are the assessment approaches used by English language teachers in Saudi universities?

To answer the first research question, the analysis included descriptive statistics and exploratory factor analysis using principal axis factoring with varimax rotation. Table 3 provides descriptive data of participant responses to part two of the survey which was concerned with the participants' assessment approaches (25 items). The items with the

**Table 3** Item level descriptive statistics

Item	<i>n</i>	Mean	SD
1. I use student assessment data to inform instructional planning and next steps for individual students and the whole class.	223	3.96	1.017
2. I monitor and revise my assessment approaches regularly.	226	4.05	0.967
3. I use a variety of formative assessment techniques (e.g., structured Q&A, feedback) and instruments (e.g., paper-pencil quizzes, personal-response systems) to check for understanding during instruction.	227	4.15	0.999
4. My summative assessment (e.g., quizzes) grades meaningfully represent individual student learning as related to curriculum expectations.	225	3.96	0.901
5. I use various summative assessment types, such as multiple-choice tests, essays, and performance-based assessments.	220	3.99	1.068
6. I engage students in monitoring their own learning and using assessment information to develop their learning skills.	224	3.99	1.037
7. I spend adequate time ensuring my assessments are responsive to and respectful of the cultural and linguistic diversity of my students.	221	3.94	0.991
8. I regularly involve students in assessment approaches during teaching.	225	3.74	1.068
9. I clearly communicate the purposes and uses of assessment to students.	228	4.03	0.926
10. I provide timely feedback to students to improve their learning.	228	4.17	0.944
11. The intended purposes of the assessment or/and the curriculum expectations influence my assessment-related decisions.	225	3.90	0.93
12. I am able to analyze and make instructional decisions based on my students' performance on unified assessments.	226	3.95	0.974
13. I monitor and revise my assessment approaches to improve the quality of my teaching.	125	3.54	0.894
14. My methods of assessment allow students to show their learning in diverse ways.	147	3.59	0.739
15. I spend adequate time differentiating my assessment approaches to meet students' specific educational needs.	144	3.44	0.859
16. I provide adequate resources and time to prepare students with special needs for assessment.	135	3.42	0.796
17. In my class, all students complete the same assignments, quizzes, and tests.	136	3.32	0.994
18. When grading student work, I use the same rubric or scoring guide for all my students.	107	3.57	0.825
19. I link my assessment tasks to learning objectives.	116	3.64	0.806
20. I am confident that my students' performance on my assessments represents what I want them to learn.	150	3.50	0.825
21. I can select assessment types that align with my learning objectives from test banks, textbook series, and/or online teacher sharing sites.	142	3.53	0.856
22. I use multiple assessment types to measure each learning objective and am confident in the grades I assign.	151	3.53	0.807
23. My students' grades and feedback are based on the data collected about their learning achievement.	153	3.56	0.768
24. I am confident that I apply my scoring guides/rubrics consistently.	143	3.57	0.756
25. I use student assessment data to inform teaching planning and next steps for individual students and the class as a whole.	156	3.53	0.791



highest means were 2, 3, 9, and 10: 2) 'I monitor and revise my assessment approaches regularly' (mean = 4.05); 3) 'I use a variety of formative assessment techniques (e.g., structured Q&A, feedback) and instruments (e.g., paper-pencil quizzes, personal-response systems) to check for understanding during instruction' (mean = 4.15); 9) 'I clearly communicate the purposes and uses of assessment to students' (mean = 4.03); and 10) 'I provide timely feedback to students to improve their learning.' (mean = 4.17). Items 15, 16, and 17, which focused on fairness (standardization, differentiations, equity), had the lowest means: 15) 'I spend adequate time differentiating my assessment approaches to meet students' specific educational needs' (mean = 3.44); 16) 'I provide adequate resources and time to prepare students with special needs for assessment' (mean = 3.42); and 17) 'In my class, all students complete the same assignments, quizzes, and tests' (mean = 3.32).

These findings suggest that the participants valued regular evaluation of their assessment approaches and used a variety of formative assessment techniques. They highly valued providing feedback to help students improve through communicating assessment purposes and uses. The items with the lowest means were concerned with assessment fairness, differentiation in assessment, and individual differences.

SPSS software (version 26) was used for data analysis. Kaiser-Meyer-Olkin (KMO) and Bartlett's test were used to evaluate the sampling adequacy for factor analysis, and to check for redundancy between variables that can be summarized with some factors. Kaiser-Meyer-Olkin Measure of Sampling Adequacy (.880) indicated that exploratory factor analysis would aid data interpretation. In addition, Bartlett's test of sphericity was significant, ( $p < .001$ ) indicating that exploratory factor analysis would aid data interpretation. Exploratory factor analysis was performed using principal axis factor with varimax rotation (Table 4). Simple factor structure (i.e., each item loading onto one factor) was sought; however, it could not be achieved with this data set. The factor loadings showed that factor 1 appeared to focus on items related to *approaches to assessment purpose and process* and included the following survey items: 1, 2, 3, and 6–11. This factor had a Cronbach's alpha (measure of internal consistency) value of 0.954. Factor 2 appeared to focus on items related to *approaches to assessment fairness and theory and* included the following survey items: 15–25. This factor had a Cronbach's alpha value of 0.959. Factor 3 appeared to focus on items related to *the use of assessment data* and included the following survey items: 4, 12, 13, and 14. This factor had a Cronbach's alpha value of 0.822. The results suggested that the sample had similar conceptions regarding assessment and similar patterns of assessment approaches. The participants highly endorsed assessment purposes (formative and summative assessment) and assessment processes (design and communication) and focused less on fairness and measurement theory (validity, reliability, mixed).

#### **RQ2: What are English teaching staff assessment training needs and preferred methods of training?**

The teachers were asked to respond regarding their current needs for professional assessment development. The participants mentioned feedback ( $N = 57$ ), peer-assessment ( $N = 55$ ), writing test items ( $N = 45$ ), and marking and scoring ( $N = 41$ ) as their top current needs for assessment education. Other participants ( $N = 23$ ) cited that they



**Table 4** Factor analysis: factor loadings

Item	Factor 1	Factor 2	Factor 3
1. I use student assessment data to inform instructional planning and next steps for individual students and the whole class.	0.531		
2. I monitor and revise my assessment approaches regularly.	0.613		
3. I use a variety of formative assessment techniques (e.g., structured Q&A, feedback) and instruments (e.g., paper-pencil quizzes, personal-response systems) to check for understanding during instruction.	0.543		
4. My summative assessment (e.g., quizzes) grades meaningfully represent individual student learning as related to curriculum expectations.			0.477
5. I use various summative assessment types, such as multiple-choice tests, essays, and performance-based assessments.	0.565		
6. I engage students in monitoring their own learning and using assessment information to develop their learning skills.	0.668		
7. I spend adequate time ensuring my assessments are responsive to and respectful of the cultural and linguistic diversity of my students.	0.584		
8. I regularly involve students in assessment approaches during teaching.	0.617		
9. I clearly communicate the purposes and uses of assessment to students.	0.589		
10. I provide timely feedback to students to improve their learning.	0.616		
11. The intended purposes of the assessment or/and the curriculum expectations influence my assessment-related decisions.	0.421		
12. I am able to analyse and make instructional decisions based on my students' performance on unified assessments.			0.429
13. I monitor and revise my assessment approaches to improve the quality of my teaching.	0.546		0.607
14. My methods of assessment allow students to show their learning in diverse ways.			0.684
15. I spend adequate time differentiating my assessment approaches to meet students' specific educational needs.	0.417	0.548	
16. I provide adequate resources and time to prepare students with special needs for assessment.	0.556	0.568	
17. In my class, all students complete the same assignments, quizzes, and tests.		0.618	
18. When grading student work, I use the same rubric or scoring guide for all my students.		0.597	
19. I link my assessment tasks to learning objectives.		0.723	
20. I am confident that my students' performance on my assessments represents what I want them to learn.		0.613	
21. I can select assessment types that align with my learning objectives from test banks, textbook series, and/or online teacher sharing sites.	0.47	0.522	
22. I use multiple assessment types to measure each learning objective and am confident in the grades I assign.		0.465	
23. My students' grades and feedback are based on the data collected about their learning achievement.	0.406	0.494	
24. I am confident that I apply my scoring guides/rubrics consistently.		0.591	
25. I use student assessment data to inform teaching planning and next steps for individual students and the class as a whole.		0.529	0.465
Factor score (Mean [SD])	3.91 (.64)	3.55 (.55)	3.82 (.75)
Percent variance explained	21.99	18.11	11.45
Cronbach's alpha	.954	.959	.822

needed to learn more about assessment in general and various assessment approaches and techniques (Table 5).

### Preferred training methods

Six items in the third part focused on preferred methods of professional learning. In the first section, the participants were asked to choose their preferred assessment methods (Table 6). The results are presented in terms of frequency counts. A total of 56 different combinations of assessment methods selected by participants were examined and no clear preference of method was found. The overall result showed that the participants wanted to learn about assessment in various ways rather than a specific method. Some methods were chosen more than others, such as independent study ( $N = 102$ ), university-based professional development sessions ( $N = 98$ ), and conferences, seminars, and workshops ( $N = 101$ ). The least favourite was attending a course in another university ( $N = 25$ ).

This finding suggests that the participants did not agree on the preferred methods for learning about assessment, indicating that teachers require diverse of ways in which professional support is provided to learn about assessment. Therefore, institutions should provide and support various professional development options to enhance teachers' assessment literacy.

### RQ3: What is the impact of the demographic characteristics of English teaching staff on their assessment approaches and professional development in assessment needs?

To identify the impact of demographic characteristics, such as gender, experience in the professional role, and assessment education, of the participants on their assessment approaches, two methods were used: t-test and analysis of variance (ANOVA) with demographic variables as random factors and factor scores as the dependent variables and chi-square tests with crosstabulation tables.

Gender was found to be statistically significant with factor 1 (assessment purposes and processes). Equal variances for each factor could not be assumed, as Levene's test was significant. Women endorsed factor 1 statistically significantly more than men. Cohen's  $d$  for this difference was 0.34. These results suggested that women valued assessment purposes and assessment processes more than men ( $t = 2.697$ ;  $df = 228$ ; Sig. [2-tailed] = 0.008; Table 7). No statistical differences were identified based on education level, job title, degree (field), or career stage of participants, indicating that the participants were similar in their assessment approach perceptions.

**Table 5** Assessment training needs

Area of training needs	N. of participants
1. Peer assessment	55
2. Feedback	57
3. Tests items writing	45
4. Marking and scoring	41
5. Test items analyses	2
6. General training	23
7. Online assessment and scoring, alternative methods of assessment	4

**Table 6** Frequency counts for preferred assessment methods

Method	Frequency
Independent study, such as professional reading, including books or online sources	102
With a peer mentor or coach	61
Small-group learning, such as professional learning community	86
University-based professional development sessions	98
Through another university course	25
Through a professional conference, seminar, or workshop	101
Through online learning	76

Experience in a professional role (novice, competent, expert) was significant in relation to factor 2 (assessment fairness and theory). Bonferonni post hoc test was performed for factor 2 and was significant at the 0.05 level. No differences between groups were found for factors 1 or 3. As shown in Table 8, a statistically significant difference was noted for factor 2 with a mean difference of  $-.35470$  (significant at the 0.05 level). Bonferonni post hoc analysis revealed that novice participants endorsed factor 2 at a significantly lower level than competent participants. Equal variances could be assumed for all factors. No statistically significant differences were noted between participants who identified themselves as decision-makers regarding assessment and those who were not decision-makers, and between those who took a course in assessment and those who took no courses.

**RQ4: What is the impact of the demographic characteristics of English teaching staff on their professional development in assessment needs?**

The relationship between preferred methods of assessment education and demographic variables was explored using the chi-square test. Three significant relationships were identified: (a) gender and learning assessment with a peer/coach; (b) participation in a course in assessment and learning assessment independently; and (c) participation in a course in assessment and university-based professional development. The data analysis showed statistically significant correlations between preferred assessment education methods and demographic variables. Women were statistically less likely to learn with a peer/coach than men. A total of 33 women selected learning from a peer/coach, while the expected number was 40.9. The opposite trend was noted for men, with 28 men selecting learning from a peer/coach when only 20.1 were expected. Pearson chi-square

**Table 7** Significance table for assessment approaches by gender

	Levene’s test for equality of variances		t test for equality of means			
	F	Sig.	t	df	Sig. (2-tailed)	
FAC 1	Equal variances assumed	13.881	0.000	2.697	228	0.008
	Equal variances not assumed			2.319	107.695	0.022
FAC 2	Equal variances assumed	10.923	0.001	2.201	199	0.029
	Equal variances not assumed			1.879	93.956	0.063
FAC 3	Equal variances assumed	8.599	0.004	1.99	227	0.048
	Equal variances not assumed			1.758	113.31	0.081

**Table 8** Bonferonni post hoc analysis for factor 2

Dependent Variable	(I) Role_ Num	(J) Role_ Num	Mean difference (I-J)	Std. error	Sig.	95% confidence interval	
Factor2	Novice	Competent	-.35470 <sup>a</sup>	0.12567	0.016	-0.6581	0.0513
		Expert	-0.29117	0.12772	0.071	-0.5996	0.0172
	Competent	Novice	.35470 <sup>a</sup>	0.12567	0.016	0.0513	0.6581
		Expert	0.06353	0.0814	1	-0.133	0.2601
	Expert	Novice	0.29117	0.12772	0.071	-0.0172	0.5996
		Competent	-0.06353	0.0814	1	-0.2601	0.133

<sup>a</sup>The mean difference was significant at the 0.05 level

statistics were reported. The *p* value for this test was 0.015, lower than an alpha value of 0.05 (a commonly used value); therefore, the findings were interpreted as significant. The same pattern was observed for the two other significant relationships reported. Participants who took a course in assessment were more likely to prefer to learn independently and through university-based professional development courses.

**Discussion**

This study explores assessment approaches, assessment training needs, and preferred methods for training English language teachers in Saudi language centres and institutes. Given the lack of assessment standards in MENA, utilising the existing measures to explore teachers’ approaches provided an insight into how their approaches were informed by contemporary assessment standards constructed for English speaking countries. The results suggested that the participants used similar methods for endorsing and valuing assessments. They reported similar training needs in various aspects of assessment while there was a variation in the preferred training methods. Even though the participants’ demographics were varied, they taught the same courses under the same policy in their institutions, which explains why their endorsements were similar. English teachers’ roles in assessment in HE remains limited given that assessment is unified, and teachers teach the subject and assess students on limited tasks that carry a small weight on the students’ overall grade. Additionally, their institutions remained similar and limited.

When it comes to teaching, learning, and assessment standards in Saudi Arabia, the National Commission for Academic Accreditation and Assessment (NCAAA) is committed to teaching with various published documents and workshops supported by deanships of quality assurance in Saudi universities (NCAAA, 2015). However, assessments seem to be left out from these documents and workshops (Almossa, 2018; 2021). As a result of a lack of assessment education, specific training programs, and standards, there is a likelihood that, for their assessment practices, English teachers in Saudi HE rely on personal beliefs experiences, institution assessment culture, and PD opportunities. Almossa (2021) reported that English teachers reported that they did not have equal access to paid PD opportunities as several factors interfered, such as nationality, family situation, center/institute policy, and university fund policies.

The research findings were in line with Coombs et al. (2018), who explored teachers’ approaches to assessment in relation to career stage. Stable approaches to assessment were reported. The findings suggested that while teachers appeared to have similar

assessment approaches, they may operationalize these approaches differently. For instance, teachers who lean toward using formative assessment may hold a different conceptualization and implementation. Almoosa (2018) observed teachers in a Saudi university English language centre and noted that, although teachers reported using a variety of formative assessment techniques, in-classroom observation did not support this, which indicated that classroom reality differed from the understanding of formative assessment. In his work, Hakim (2015) reported a mismatch between learning outcomes and observed teachers' practices and how they perceived their work to be aligned with the learning outcomes. Therefore, what might be reflected as a shared understating of assessment might not be the reality inside the classroom or during the assessment process. Therefore, knowledge of current thinking about assessment practices and education can help to shape assessment literacy development priorities.

The assessment PD needs of the teachers revealed that their top priorities were learning more about using peer assessments and feedback, which fall into the formative assessment category. Additionally, some teachers wanted to learn more about writing test items and marking and scoring, which are important parts of the summative assessment process. The teachers wanted to balance between boarding their knowledge about summative and formative assessment techniques. These findings are in line with Almoosa (2021) and Almoosa (2018). Ezza (2017) suggested training programs on different aspects of assessment design to be provided to university teachers that are formal. This echoes what the teachers in current study teachers' want to focus on developing their assessment design.

While teachers were similar in terms of education needs, preferred learning methods were different, it is important that policy-makers consider differentiated assessment education and professional development opportunities (Deluca et al. 2018).

## Conclusion

This study provided empirical evidence on approaches to assessment purposes, assessment processes, assessment fairness, and measurement theory. The need for this investigation comes from the importance of exploring teachers' practices to understand how they approach assessment with their own conceptualization of it. Assessment literacy and assessment practices deserve policy-makers' attention as improving the quality of learning outcomes is a huge project in Saudi higher education. There are several research implications to be drawn from the study. First, it is highly recommended that NCAAA design and publish an assessment standards booklet that elaborates upon assessment principles and expectations. Second, designing PD programs should be built in response to teachers' specific needs not the other way around to suit teachers' learning needs and preferences. Third, a variety of assessment literacy opportunities should be provided to promote self-study methods. Offering a wide range of books or eBooks and journals, granting teachers the access to online learning resources, webinars, organizations, hands-on websites that offer practical tips, and assessment models through institutional subscription or allow borrowing from the centre library would go a long way in contributing to their learning.

This study faces several limitations. First, the sampling size was small given the teachers' response rate; thousands of teachers were invited but only a few responded. Second, the study relied on self-report from a self-reporting instrument in a specific

context. Data triangulation could provide further explanations for results. Third, a modified version of ACAI was used to contextualize the instrument to the targeted population. Consequently, the scenarios section in the survey was not adopted based on the reviews received during the pilot study.

Further research exploring assessment literacy in different contexts using contemporary measurements is required. Studies should explore differences in assessment perception and approaches among teachers from various contexts and backgrounds to provide an understanding of assessment literacy that considers teachers' knowledge and contexts (Willis et al. 2013). Also, future research should explore gender variables in a segregated system, as some teachers were educated in gender-segregated institutions.

#### Abbreviations

ACAI: Approaches to Classroom Assessment Inventory; ANOVA: Analysis of variance; MENA: Middle East and North Africa; PD: Professional development; NCAAA: National Commission for Academic Accreditation and Assessment

#### Acknowledgements

We would like to thank all the participants for their generous time. A special thanks to Dr. Christopher Deluca for his support during the design and data analysis of the study and Mrs. Reem Alsayegh and Dr. Suhad Sonbul for their contribution to the early versions of the adopted survey.

#### Code availability

Not applicable.

#### Authors' contributions

SA conceived and planned the research design and data collection procedures. SA and SH contributed to the data collection process, the implementation of the research, and to the analysis of the results. SA and SH wrote the manuscript. Both authors read and approved the final manuscript.

#### Authors' information

Samar Almoosa is an Assistant Professor in Assessment at Umm Al-Qura University in Saudi Arabia. She holds a PhD in Applied Linguistics from King's College London. Her research interests are focused around learning and teaching experiences in higher education context, assessment practices and assessment literacy.

Sahar M. Alzahrani is an Assistant Professor of Applied Linguistics, English Language Centre, Umm Al-Qura University. Her PhD is from University of Southampton, UK. Interests: learner autonomy, innovation in teaching, blended learning, MOOCs, and online instructional design. Awards: LASIG Antwerp best poster; University of Southampton 3MT competition; and CALICO Outstanding Graduate Student Award.

#### Funding

This work was supported by the Deanship of Scientific Research at Umm Al-Qura University for supporting this work under Grant Code: 19-EDU-1-02-0005.

#### Availability of data and materials

Available upon request.

#### Declaration

#### Competing interests

The authors declare that they have no competing interests.

Received: 21 January 2022 Accepted: 18 March 2022

Published online: 12 April 2022

#### References

- Almansory, M. (2016). *EFL teachers' beliefs and attitudes towards English language assessment in a Saudi university's English Language Institute (Unpublished doctoral dissertation)*. Exeter: University of Exeter.
- Almoosa, S. (2018). *Developing pedagogy and assessment in EFL: A case study of a Saudi university*. [Unpublished doctoral dissertation]. London: King's College London.
- Almoosa, S. (2021). Assessment literacy: transnational teachers' concepts, practices, and professional development needs. In A. Ahmed, & O. Barnawi (Eds.), *Mobility of knowledge, practice and pedagogy in TESOL teacher education: Implications for transnational contexts* (pp. 83–106). Springer International Publishing. [https://doi.org/10.1007/978-3-030-64140-5\\_5](https://doi.org/10.1007/978-3-030-64140-5_5).
- Barnes, N., Fives, H., & Dacey, C. M. (2017). US teachers' conceptions of the purposes of assessment. *Teaching and Teacher Education*, 65, 107–116.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: principles, policy & practice*, 5(1), 7–74.
- Brookhart, S. M. (2011). Educational assessment knowledge and skills for teachers. *Educational Measurement: Issues and Practice*, 30(1), 3–12.

- Brown, G. T. (2004). Teachers' conceptions of assessment: Implications for policy and professional development. *Assessment in Education: Principles, Policy & Practice*, 11(3), 301–318.
- Cauley, K. M., & McMillan, J. H. (2010). Formative assessment techniques to support student motivation and achievement. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, 83(1), 1–6. <https://doi.org/10.1080/00098650903267784>.
- Cizek, G. J., Fitzgerald, S. M., & Rachor, R. A. (1995). Teachers' assessment practices: Preparation, isolation, and the kitchen sink. *Educational Assessment*, 3(2), 159–179.
- Coombs, A., DeLuca, C., LaPointe-McEwan, D., & Chalas, A. (2018). Changing approaches to classroom assessment: An empirical study across teacher career stages. *Teaching and Teacher Education*, 71, 134–144.
- Deluca, C., LaPointe-McEwan, D., & Luhanga, U. (2016a). Teacher assessment literacy: a review of international standards and measures. *Educational Assessment, Evaluation and Accountability*, 28(3), 251–272.
- DeLuca, C., LaPointe-McEwan, D., & Luhanga, U. (2016b). Approaches to classroom assessment inventory: A new instrument to support teacher assessment literacy. *Educational Assessment*, 21(4), 248–266. <https://doi.org/10.1080/10627197.2016.1236677>.
- DeLuca, C., Valiquette, A., Coombs, A., LaPointe-McEwan, D., & Luhanga, U. (2018). Teachers' approaches to classroom assessment: A large-scale survey. *Assessment in Education: Principles, Policy & Practice*, 25(4), 355–375.
- Ezza, E. S. (2017). Criteria for assessing writing at Majmah University. In S. Hidri, & C. Coombe (Eds.), *Evaluation in foreign language education in the Middle East and North Africa, second language learning and teaching*, (pp. 185–200). Springer International Publishing.
- Gotch, C. M., & French, B. F. (2014). A systematic review of assessment literacy measures. *Educational Measurement: Issues and Practice*, 33(2), 14–18.
- Hakim, B. (2015). English language teachers' ideology of ELT assessment literacy. *International Journal of Education and Literacy Studies*, 14(1), 75–97.
- Harlen, W. (2006). On the relationship between assessment for formative and summative purposes. *Assessment and Learning*, 2, 95–110.
- Hattie, J. (2008). *Visible learning: a synthesis of over 800 meta-analyses relating to achievement*. Routledge.
- Herppich, S., Praetorius, A.-K., Förster, N., Glogger-Frey, I., Karst, K., Leutner, D., ... Klug, J. (2018). Teachers' assessment competence: Integrating knowledge-, process-, and product-oriented approaches into a competence-oriented conceptual model. *Teaching and Teacher Education*, 76, 181–193.
- Looney, A., Cumming, J., van Der Kleij, F., & Harris, K. (2018). Reconceptualising the role of teachers as assessors: teacher assessment identity. *Assessment in Education: Principles, Policy & Practice*, 25(5), 442–467.
- NCAAA (2015). *Handbook for quality assurance and accreditation in Saudi Arabia part 2: Internal quality assurance arrangements*. NCAAA <http://www.kfupm.edu.sa/deanships/dad/Documents/AAC/NCAAA%20Documents/H2.%20Handbook%20Part%202.pdf>. Accessed 8 Jan 2022.
- Popham, W. J. (2013). Can classroom assessments of student growth be credibly used to evaluate teachers? *English Journal*, 103(1), 34–39.
- Rauf, M., & McCallum, L. (2020). Language assessment literacy: Task analysis in Saudi universities. In *The Assessment of L2 Written English across the MENA Region* (pp. 13–41). Cham: Palgrave Macmillan.
- Umer, M., Zakaria, M. H., & Alshara, M. A. (2018). Investigating Saudi University EFL teachers' assessment literacy: theory and practice. *International Journal of English Linguistics*, 8(3), 345–356.
- Willis, J., Adie, L., & Klenowski, V. (2013). Conceptualising teachers' assessment literacies in an era of curriculum and assessment reform. *The Australian Educational Researcher*, 40(2), 241–256. <https://doi.org/10.1007/s13384-013-0089-9>.
- Xu, Y., & Brown, G. T. (2016). Teacher assessment literacy in practice: A reconceptualization. *Teaching and Teacher Education*, 58, 149–162.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

---

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)

---