Journal of English for Specific Purposes Praxis

Journal Site: https://www.jespp.ir ©2024 - Volume: 1, Issue: 1, pp. 41–64

Article Citation: ISSN: 3092-7854

Kousha, M., Ghalibafan, M. S., & Souri, M. (2024). The status quo of English language assessment literacy among Iranian ESP teachers of computer sciences engineering. *Journal of English for Specific Purposes Praxis, 1*(1), 41–64. https://doi.org/10.22034/jespp.2024.217546

JESPP

The Status Quo of English Language Assessment Literacy Among Iranian ESP Teachers of Computer Sciences Engineering

Mahdi Kousha[1], Gholam-Reza Abbasian[2], Mohammad Sadegh Ghalibafan[3], and Mehrdad Souri[4]

[1] Department of English Language, Faculty of Foreign Languages, South Tehran Branch, Islamic Azad University, Tehran, Iran; m_koosha2011@yahoo.com

[2] English Language Department, Faculty of Basic Sciences, Imam Ali University, Tehran, Iran; gabbasian@gmail.com

[3] Corresponding author, Department of English Language, Faculty of Foreign Languages, South Tehran Branch, Islamic Azad University, Tehran, Iran; msghalibafan@gmail.com

[4] Department of English Language, Faculty of Foreign Languages, South Tehran Branch, Islamic Azad University, Tehran, Iran; mehrdadsouri1991@gmail.com

Original Research Article | Date of Submission: 05 April 2024 | Date of Acceptance: 28 June 2024

Abstract

Assessment is a vital aspect of the education process, and investigators have placed it at the heart of the teaching and learning process. Teachers spend a significant amount of time engaging in assessment activities to promote learning; thus, it is of paramount importance that teachers have an adequate level of assessment literacy. To shed more lights on this issue in the context of Iranian universities, this study intended to measure and describe the relative level of assessment literacy of Iranian ESP teachers of computer sciences engineering and to define the rank of each sub-scale of assessment literacy among them. This study adopted a quantitative research design and collected data from 60 Iranian ESP teachers of computer sciences engineering by responding to the "Classroom Assessment Knowledge (CAK)" scale. Analyses of the data revealed that first, Iranian teachers of computer sciences engineering did not have acceptable level of assessment literacy; second, there were significant differences among the nine sub-scales of assessment literacy; third, the ESP teachers suffered from significant deficiencies which not only indicate weaknesses in teacher education programs as far as language testing and assessment are concerned but also are persuasive enough to render specific assessment education for each group involved in foreign language teaching ESP programs. Findings can be helpful for ESP teachers, educators, assessors, and program developers.

Keywords: assessment literacy, computer sciences engineering, ESP assessment, Iranian ESP teachers

1. Introduction

Assessment refers to "the process of obtaining information that is used to make educational decisions about students, to give feedback to the student about his/her progress, strengths, and weaknesses, to judge instructional effectiveness and curricular adequacy, and to inform policy" (AFT, NCME, & NEA, 1990; as cited in Davidheiser, 2013, p. 10). Assessment is also a central part of

language teaching and learning since the information collected through assessment gives idea about the quality of instruction and development of students concerning having adequate knowledge and skills in language. In other words, it is interwoven with teaching and learning, and involves making judgments about learners and monitoring their development in order to assess their needs and tailor instruction to optimize learning (Gkogkou & Kofou, 2021). In this sense, teachers have critical roles for making assessment process effective and useful. They use a variety of assessment tools in language classes which help them observe the strengths and weaknesses of teaching and learning process, and also provide evidence for re-designing their instructions (Çakır & Genç, 2022).

According to Giraldo (2018), the first suggestion to assessment literacy in education was introduced in their Standards for Teacher Competence in Educational Assessment of Students by the American Federation of Teachers, National Council on Measurement in Education, and National Education Association (1990). In fact, they thought these guidelines were needed to help teachers become aware of assessment in and out of classroom contexts. The guidelines can be classified into two strands. The first one relates to instruction; teachers should be able to select, design, and evaluate valid assessment for positive effects on learning, teaching, and schools. Uses of tests and test results was the second one strand; teachers are expected to know when assessments are being used improperly, and to know how to communicate results well to different stakeholders. This shows that assessment plays an important role in many people's lives (McNamara, 2000), and teachers therefore need to be "competent in the principles and practice of language assessment" (Harding & Kremmel, 2016).

Later, Richard Stiggins (1991) used the term assessment literacy to include knowledge and skills that stakeholders such as teachers and school administrators should have about assessment. He states that educators with assessment literacy know what they assess, how to assess, what the possible problems with assessment are, and how to prevent them from occurring. Popham (2009) believes that assessment literacy is "an individual's understandings of the fundamental assessment concepts and procedures deemed likely to influence educational decisions". He also states that "assessment literacy goes well beyond teachers possessing knowledge of how to interpret assessment data but also includes in-depth knowledge of issues such as test reliability and test validity" (p. 267). Popham (2009) continues to say that assessment literacy entails knowledge of reliability and threats to it, tests' content validity, fairness, use of alternative assessments such as portfolios, formative assessment, student test design of closed-ended and open-ended test tasks preparation, as well as assessment of ESP learners.

In history, assessment literacy has extended teachers' toolbox to monitor, record, improve, and report on student learning. In fact, there has also been increasing attention as to how assessment has consequences on teaching, learning, and school curricula (Popham, 2009, 2011); this attention has led to a belief that teachers should have a critical stance towards how assessment impacts stakeholders. There has been growing interest in examining teachers' mastery and application of assessment literacy in recent years, which addresses the knowledge and skills that teachers need in order to use assessment effectively to evaluate and expand students' learning (Davidheiser, 2013; Howell, 2013; Mertler, 2003; Plake & Impara, 1993; Stiggins, 1995, 1999). This, according to Webb (2002), can be attributed to (1) the use of a variety of methods of assessment for the purpose of formative and summative assessments that is now considered to be of great significance, and (2) the development of student-centered learning. Furthermore, students have increased the need and expectation on teachers to understand how students' learning can be adequately assessed, to determine whether students have achieved the learning objectives and what meaning should be given to the assessment information gathered.

Elshawa et al. (2016) have identified different key elements that should be taken into account once aiming to improve learning through assessment: first, communication of constructive feedback to students; second, having students involved in learning process; third, analysis and consideration of the assessment results; fourth, pinpointing the important effects of assessment on the motivation and self-esteem of students; and finally, helping students improve by keeping them involved in assessing themselves.

2. Literature Review

2.1. Assessment Literacy

Assessment literacy refers to the way of assessing students and what they know and what they can do Ukrayinska (2024). It also refers to interpreting the results and consequences of what they know, therefore, applying these results will improve the learners' learning and effectiveness of their program (Webb, 2002). The term Assessment Literacy was coined for the first time by Richard Stiggins in 1991. He contends that "educators 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". Furthermore, Stiggins (1999, 2006) argued that assessment literacy was an important factor to the quality of instruction and teachers used it to show that students understand concepts and instruction. Similarly, Bayat and Rezaei (2015) quote from Paterno (2001) saying that it calls for teachers to "possess the knowledge about the basic principles of sound assessment practice, including terminology, the development and use of assessment methodologies and techniques, familiarity with standards of quality in assessment and familiarity with alternative to traditional measurements of learning" (p. 140). According to Elshawa et al. (2016, p. 135), assessment literacy can be defined as "the knowledge about how to assess what students know and can do, interpret the results of these assessments, and apply them to improve student learning".

Language assessment literacy helps teachers to know more about the learners' weaknesses and strengths as they learn a new language (Estaji et al., 2024). In traditional assessment the focus was just on the results and final consequences of the assessment (Brown, 2004). These results sometime were compared with other learners' result which is known as NRT to rank the students or compared with a criterion by the learners that should be achieved that is known as CRT (Alderson, 2000), but in assessment the main goal is to see if learning has occurred or not and it is the systematic process of gathering information for making decision about people Förster (2023). So, there was a move from assessment of learning to assessment for learning recently (Heaton, 2011).

As mentioned earlier, scholars have defined various descriptions for assessment literacy but all definitions vary depending on the context of use, but what is common in all these visions are that teachers must recognize different purposes of assessment and use them accordingly (Volante & Fazio, 2007; as cited in Bayat & Rezaei, 2015).

2.2. Literacy Assessment: Definitions, Knowledge, Principles, and Skills

Literacy assessment, according to the National Council of Teachers of English (2018), is an essential part of literacy teaching and learning and it simplifies the conditions for literacy teaching and learning. In literacy assessment, the main component of a literacy teacher's progress and practice is professional knowledge. Some of the principles in language assessment literacy are as follow:

- Literacy assessment is not a technical activity it is a social process, and the students should be assessed by knowledgeable people.
- Literacy assessments is a classroom learning context; and the students themselves help the assessors to find their strengths and weaknesses
- Literacy assessment is meaningful to the learner.
- Literacy assessment includes cognitive activities.
- Literacy assessments are valid only to the extent that they help students learn.
- Literacy assessment is purposeful, for each individual assessment.
- Literacy assessment has diversity in languages, in learning styles, and in rating and way of learning.
- Literacy assessment includes varieties of multiple measures of different domains, such as processes, texts, and reflection.

Assessment literacy is using multiple sources of information to understand students' learning and adjusting instruction based on their needs and acting based on their levels and use summative assessment for learning (National Council of Teachers of English, 2018). According to Davies (2008), there are three components in assessment literacy: knowledge, skills and principles. Knowledge refers to awareness of applied linguistics, awareness of theory and concepts, awareness of own language assessment context. Knowledge of language and language methodologies include knowledge of multilingual learners and content-based language teaching. Skills refer to instructional skills, design skills for language assessments, skills in educational measurement, and technological skills. Skills and principles are related to assessing language for designing the test and interpreting its statistics and assessing language skills ability. Principles refer to a wellness of and actions towards critical issues in language assessment. Principles are ethics, fairness, and consequences of assessment.

Based on Scarino (2013) the ingredient of language assessment literacy of language teachers is knowing about social role of assessment and political and social forces that involved this power and its consequences. He states that knowledge, skills and principles in language assessment literacy include teachers' interpretive framework. It refers to discussion; in contacts, practices, beliefs, attitudes, and theories are the things that shape language assessment literacies. Then teachers' ways of thinking are another important factor in language assessment literacy. Skills include instructional skills and the construction of the four language skills. Language assessment principles put emphasis on the presence of ethics, consequential and power of the test that the teacher should be aware of them (Shohamy, 2001).

2.3. Teachers' Assessment Literacy

The National Education Association (2003) reported that every teacher must understand the principles of sound assessment and must be able to utilize those principles as a matter of routine in doing their work for assessing student's achievement. Many surveys have studied the outcome of teacher training and professional development programs in terms of assessment literacy opposed to others who have examined assessment literacy as a student-related element such as student achievement (Howell, 2013). In the recent years, there is no doubt that teacher assessment literacy is growing in importance, in improving teaching and learning, and it moves from periphery to center stage around the world. Therefore, the need for adequate assessment training among teachers in selecting, administering, interpreting and decision making is tangible. Although teacher assessment literacy is a key factor in the success of teaching, the development of teacher's assessment literacy is also important for the

development of quality of learning and instruction, some studies concluded that teachers do not benefit from it.

These theoretical views of assessment literacy provide a starting point for understanding the importance of teacher assessment literacy. Teachers need to understand the importance of student assessment and the need for promoting the assessment competencies to gather accurate information about student performance, provide meaningful feedback to them, use assessment results effectively to increase students' achievement and this requires that they have an understanding of types of assessment and the purpose of assessment in the classroom.

2.4. Teaching Experience and Critical Cultural Awareness

Assessment is a main part of education in each educational system, although the level of assessment literacy of Iranian teachers is a controversial matter (Jan-nesar et al., 2020). According to Stiggins (1988), "teachers spend one-third of their time on assessment related activities" and it is one of the most important responsibilities of teachers to assess their students' performance. He also states that teachers are neither trained nor prepared to perform their role in the assessment process and about 47% of teachers reported their assessment training was not adequate and they are not prepared sufficiently in the area of assessment, and most of their assessment knowledge was a result of trial and error.

Additionally, Stiggins (1995) states that recent studies show that there is great inadequacy in teachers' training in terms of development, administration and interpretation of results derived from different assessment types and few teachers are prepared to face the challenges imposed on them and their own classroom through assessment. In fact, the opportunity to learn how to deal with such challenges has never been provided for teachers due to the lack of competence in assessment in teacher training programs and the inability of such programs in providing the necessary assessment knowledge for them to engage in assessing the learning occurring in the classroom.

Iranian educational setting does not seem an exception to this pattern, and assessment illiteracy of university teachers may have given learner autonomy a very far-to-reach goal, since learner autonomy has not been clearly observed as it should have been among Iranian learners. Moreover, Iranian students' lack of knowledge of their own ability to learn, may greatly affect their learning progress both inside and outside the classroom. Moreover, it seems that Iranian university teachers are no exception and there is inconsistency in the extent of assessment knowledge among them. In other words, it seems that they do not receive any training in methods of teaching including those of assessment of what they are supposed to teach. Thus, the assumption is that there is no certainty that they enjoy the acceptable level of assessment literacy.

As a result, the primary goal of this study is to define the knowledge of Iranian ESP teachers of computer sciences engineering in developing classroom assessment and to examine their level of assessment literacy related to the development of effective classroom assessment models. Specifically, the main purpose of this study was to measure and describe the relative level of assessment literacy of Iranian ESP teachers of computer sciences engineering. This problem motivated raising the following research questions.

Jan-nesar et al. (2020) have reviewed some of the studies done in Iran regarding assessment literacy. Some of these studies are reviewed in the Table 1 below:

Table 1

Assessment Study Reviews

Mellati & Khademi (2018)	Exploring Teachers' Assessment Literacy: Impact on Learners' Writing Achievements and Implications for Teacher Development	Teachers' assessment literacy inventory, semi-structured interview, non-participatory observation, and Writing Competence Rating Scale (WCRS).	10 teachers and 75 sophomore	Teacher-mediated classroom assessment	The learners' writing is significantly affected by the teachers' assessment literacy
Farhady & Tavassoli (2018)	Assessment Knowledge Needs of EFL Teachers	Fulcher's (2012) needs assessment questionnaire	246 male and female EFL teachers	Requires instruments or assessment	Most teachers have different priorities in assessment although, in general, they agreed on same topics as being important for assessment.
Ashraf & Zolfaghai (2018)	EFL Teachers' Assessment Literacy and Their Reflective Teaching	Assessment literacy inventory (Zolfaghari & Ashraf, 2015) and the reflective teaching questionnaire designed by Behzadpour (2007).	120 Iranian EFL teachers	Reflective teaching	A significant and positive relationship between assessment literacy and reflective teaching
Fard & Tabatabei (2018)	Investigating Assessment Literacy of EFL Teachers in Iran	Melter's (2003) Classroom Assessment Literacy Inventory	52 male and female EFL teachers in both the public and private sector	Classroom assessment literacy	Teachers have insufficient knowledge of assessment literacy.
Janatifar, Marandi, & Babaei (2018)	Iranian EFL Teachers' Language Assessment Literacy (LAL) under an Assessing Lens	Fulcher's (2012) LAL survey	280 male and female EFL teachers in 17 provinces	Hands-on experience of assessment	Iranian EFL teachers complain about lack of hands-on experience of assessment which has confined them to theoretical aspects of assessment literacy.
Zolfaghari & Ahmadi (2016)	Assessment literacy components across subject matters	Interview	32 EFL and non-EFL teachers (both male and female)	Main components of assessment literacy	General components: knowledge of assessment processes, ethical issues and understanding the learners' Specific components: the learners' belief about the teachers

Ahmadi& MirShojae (2016)	Iranian English Language Teachers' Assessment Literacy: The Case of Public School and Language Institute Teachers	Semi-structured interviews	20 EFL teachers	Reliability, validity, learners' engagement, social aspects of assessment, formative and alternative assessment	The EFL teachers were not assessment literate and lack skills in assessing reliability, validity, learners' engagement, social aspects of assessment, formative and alternative assessment
Rahmani (2016)	EFL Teacher Education in Iran: Does It Promote Trainees' Pedagogical Content Knowledge?	40 EFL teachers in two age groups	Successful Teacher Questi. (Moafian &Pishghadam (2009).	Teachers' professional success	Assessment literacy of the older and younger participants is significantly different.

3. Research Questions

This study measured and described the relative level of assessment literacy of Iranian ESP teachers of computer sciences engineering and to define the rank of each sub-scale of assessment literacy among them. Specifically, it sought to answer the following research questions:

1. Do the Iranian ESP teachers of computer sciences engineering have acceptable level of assessment literacy?
2. Do the Iranian ESP teachers of computer sciences engineering have acceptable level of subcategories of assessment literacy?

4. Method

4.1. Research Design

In this study, the researchers attempted to explore the level of assessment literacy for ESP university teachers without manipulating the independent variable. This study adopted a quantitative, single shot, non-experimental research design. In fact, the researchers did not have any control over the selection and manipulation of the variables. The dependent variable of this study was assessment literacy knowledge of ESP university teachers of computer sciences engineering.

4.2. Participants

The participants of this study included a pool of 60 Iranian ESP teachers of computer sciences engineering holding Ph.D./M.A. degrees (34 males and 26 female) who had at least 2 years' experience in teaching ESP textbooks (i.e., technical English for the students of computer sciences engineering). They were chosen from different state universities of Tehran, Iran including University of Sanati Sharif, University of Shahid Beheshti, University of Tehran, University of Amirkabir, University of Allame Tabatabaei, and University of Tarbiat Modares which work on the field of computer sciences engineering. They were chosen through convenience sampling based on the certain criteria such as geographical area, availability at a certain time, ease of accessibility and their willingness to participate in this study. The sample did not include university teachers with B.S. degree and university teachers with lower than 2 years teaching experience in their fields based on the demographic information that was collected on all participants such as gender, age, highest level of education and subject(s) taught.

4.3. Instruments

To examine and measure the assessment knowledge of teachers and to investigate the research questions, Tao's (2014) scale, Classroom Assessment Knowledge (CAK) (see Appendix) comprised of 27 multiple- choice items (three items for each standard) which are concerned with the teachers' assessment knowledge. The Cronbach's alpha reliability index of the questionnaire reached 0.78.

Concerning construct validity, the scale consists of four main factor. Its first factor i.e., "Audience Rights in Assessment" consists of three subscales of "Developing Valid Grading Procedure", "Recognizing Unethical Assessment Practices" and "Keeping Accurate Records of Assessment Information". The second factor i.e., "Assessment Process" includes two subscales of "Choosing Appropriate Assessment Methods" and "Administering, Scoring and Interpreting Assessment Results". The third factor refer to "Developing Process" consists of two subscales of "Ensuring Quality Management" and "Developing Appropriate Assessment Methods". And finally; "Using Assessment Results for Decision Making" and "Communicating Assessment Results" refer to "Application of Assessment Results" as the fourth factor.

4.4. Procedure

To perform the research, The Classroom Assessment Knowledge test was distributed to 60 (34 males and 26 females) Iranian ESP teachers of computer sciences engineering holding at least Ph.D./MA degrees who were volunteered to participate in this study for collecting quantitative data. They had at least 2 years' experience in teaching at different state universities in Tehran, Iran during academic year 2023-2024. The tests were sent to them electronically through email, Telegram, or WhatsApp. Teachers were asked to complete the test at their convenience and they were assured that the responses would be confidential. Finally, the data collected were entered into SPSS Program Software Version 24.00 to be analyzed in response to the research questions.

4.5. Statistical Analysis

One-sample t-tests were run to probe the first research question of the study that asked about the Iranian ESP teachers of computer sciences engineering acceptable level of assessment literacy. In addition, a Multivariate Analysis of Variances (MANOVA) was run to explore the second research question which asked about the Iranian ESP computer sciences engineering teachers have acceptable level of subcategories of assessment literacy.

5. Results

The data were analyzed through using parametric tests of one-sample t-test and MANOVA which has a common assumption; i.e., normality of the data. As displayed in Table 2, the ratios of skewness and kurtosis over their respective standard errors were lower than the absolute value of 1.96; hence normality of the data was assured.

Table 2

Skewness and Kurtosis Testing Normality Assumption

Group		*N*	Skewness			Kurtosis		
		Statistic	Statistic	Std. Error	Ratio	Statistic	Std. Error	Ratio
ESP	ChooseAA	60	0.196	0.374	0.52	-1.344	0.733	-1.83
	DevelopA	60	0.015	0.374	0.04	-0.492	0.733	-0.67
	AdminSI	60	0.534	0.374	1.43	-0.071	0.733	-0.10
	DevelopVGP	60	0.309	0.374	0.83	0.211	0.733	0.29
	UsingARDM	60	-0.301	0.374	-0.80	-1.114	0.733	-1.52
	RecognizingUAP	60	0.485	0.374	1.30	0.301	0.733	0.41
	KeepARA	60	0.456	0.374	1.22	-0.438	0.733	-0.60
	EnsureQM	60	0.574	0.374	1.53	-0.433	0.733	-0.59
	CommunicateAR	60	0.369	0.374	0.99	-0.758	0.733	-1.03

Note. ChooseAA = Choosing Appropriate Assessment methods, DevelopA = Developing Assessment Method, AdminSI = Administering, Scoring and Interpreting Assessment Results, DevelopVGP = Developing Valid Grading Procedure, UsingARDM = Using Assessment Results for Decision Making, RecognizingUAP = Recognizing Unethical Assessment Practices, KeepARA = Keeping Accurate Records of Assessment, EnsureQM = Ensuring Quality Management Assessment Practices and CommunicateAR.

As displayed in Table 3 and Figure 1, the Iranian ESP teachers of computer sciences engineering means on sub-scales of assessment literacy ranged from a high mean of 1.65 on "Using Assessment Results for Decision Making" to a low of one on "Recognizing Unethical Assessment Practices".

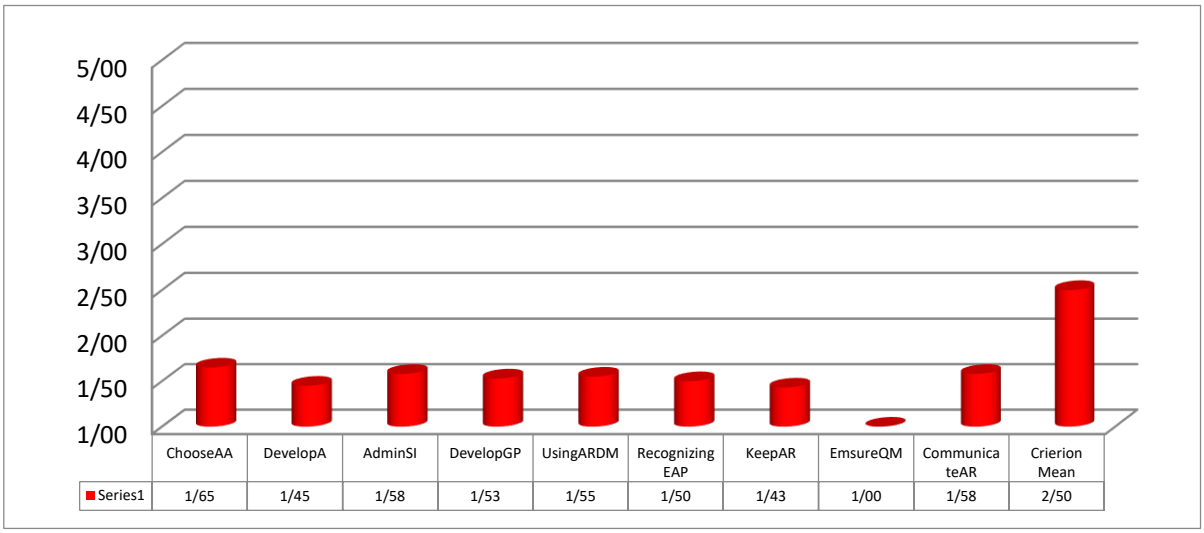
Table 3

Descriptive Statistics; Sub-Scales of Assessment Literacy for Iranian ESP Teachers of Computer Sciences Engineering

Assessment	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
ChooseAA	1.50	0.152	1.193	1.807
DevelopA	1.57	0.123	1.325	1.825
AdminSI	1.55	0.134	1.279	1.821
DevelopVGP	1.42	0.123	1.175	1.675
UsingARDM	1.65	0.154	1.338	1.962
RecognizingUAP	1.00	0.143	0.710	1.290
KeepARA	1.57	0.107	1.359	1.791
EnsureQM	1.45	0.147	1.152	1.748
CommunicateAR	1.52	0.119	1.285	1.765

Figure 1

Mean Scores on Assessment Literacy of ESP Computer Sciences Engineering Teachers



One-sample t-test was run to compare the Iranian ESP computer sciences engineering teachers' means on the nine sub-scales of assessment literacy against the criterion mean of 2.50 (based on the Likert scale that is between 1 to 5, the criterion mean was selected as 2.5). Based on the results displayed in Table 4, it can be concluded that:

1-1: The ESP computer sciences engineering teachers' mean score on sub-scale of "Choosing Appropriate Assessment Methods", was 1.5. This amount of mean was significantly lower than the criterion mean of 2.50, $t(59) = -6.583$, $p = 0.000$, $r = 0.725$, (representing a large effect size).

Table 4

One-Sample t-test; Iranian ESP teachers of Computer Sciences Engineering' Means on Sub-Scales of Assessment Literacy

	df	M	SD	SEM	T	Sig.(2-tailed)	Mean Diff.	95% CI Lower	Upper
ChooseAA	60	1.50	0.961	0.152	-6.583	0.000	-1.00	-1.31	-0.69
DevelopA	60	1.58	0.781	0.123	-7.493	0.000	-0.92	-1.17	-0.68
AdminSI	60	1.55	0.846	0.134	-7.104	0.000	-0.95	-1.22	-0.68
DevelopVGP	60	1.43	0.781	0.123	-8.708	0.000	-1.07	-1.32	-0.83
UsingARDM	60	1.65	0.975	0.154	-5.512	0.000	-0.85	-1.16	-0.54
RecognizingUAP	60	1.00	0.906	0.143	-0.473	0.000	-1.50	-1.79	-1.21
KeepARA	60	1.58	0.675	0.107	-8.666	0.000	-0.92	-1.14	-0.71
EnsureQM	60	1.45	0.932	0.147	-7.123	0.000	-1.05	-1.35	-0.75
CommunicateAR	60	1.53	0.751	0.119	-8.215	0.000	-0.97	-1.22	-0.73

1-2 The ESP computer sciences engineering teachers' mean score on sub-scale of "Developing Assessment Methods" was 1.58. This amount of mean was significantly lower than the criterion mean of 2.50 ($t(59) = -7.493$, $p = 0.000$, $r = 0.768$ representing a large effect size).

1-3 The ESP computer sciences engineering teachers' mean score on sub-scale of "Administering, Scoring and Interpreting Assessment Results" was 1.55. This amount of mean was significantly lower than the criterion mean of 2.50 ($t(59) = -7.104$, $p = 0.000$, $r = 0.751$ representing a large effect size).

1-4 The ESP computer sciences engineering teachers' mean score on sub-scale of "Developing Valid Grading Procedures" was 1.43. This amount of mean was significantly lower than the criterion mean of 2.50 ($t(59) = -8.708$, $p = 0.000$, $r = 0.812$ representing a large effect size).

1-5 The ESP computer sciences engineering teachers' mean score on sub-scale of "Using Assessment Results for Decision Making" was 1.65. This amount of mean was significantly lower than the criterion mean of 2.50 ($t(59) = -5.512$, $p = 0.000$, $r = 0.662$ representing a large effect size).

1-6 The ESP computer sciences engineering teachers' mean score on sub-scale of "Recognizing Unethical Assessment Practices" was 1. This amount of mean was significantly lower than the criterion mean of 2.50 ($t(59) = -10.473$, p =0 .000, $r = 0.859$ representing a large effect size).

1-7 The ESP computer sciences engineering teachers' mean score on sub-scale of "Keeping Accurate Records of Assessment Information" was 1.58. This amount of mean was significantly lower than the criterion mean of 2.50 ($t(59) = -8.666$, $p = 0.000$, $r = 0.811$ representing a large effect size).

1-8 The ESP computer sciences engineering teachers' mean score on sub-scale of "Ensuring Quality Management" was 1.45. This amount of mean was significantly lower than the criterion mean of 2.50 ($t(59) = -7.123$, $p = 0.000$, $r = 0.752$ representing a large effect size).

1-9 The ESP computer sciences engineering teachers' mean score on sub-scale of "Communicating Assessment Results" was 1.53. This amount of mean was significantly lower than the criterion mean of 2.50 ($t(59) = -8.215$, $p = 0.000$, $r = .0796$ representing a large effect size).

Since nine separate one-sample t-tests were run to probe the first research question, the Benjamini-Hochberg False-Discovery-Rate (FDR) should be computed to remove the possibility of inflated error rate. The FDR method requires computing corrected alpha values for each successive application of a single statistical analysis. Since all p-values mentioned above were .000; and since they were lower than the lowest corrected alpha level (i.e., 0.05 / 8 = 0.006), it can be claimed that Iranian ESP computer sciences engineering teachers did not have acceptable level of assessment literacy.

Then, in order to answer the second research question, a Multivariate Analysis of Variances (MANOVA) was run to compare the Iranian ESP teachers of computer sciences engineering means on the nine sub-scales of assessment literacy. Based on the results displayed in Table 5, $F(8, 32) = 2.96$, $p = 0.013$, Partial $\eta 2 = 0.426$, (representing a large effect size) it can be concluded that there were significant differences between the Iranian ESP teachers of computer sciences engineering means on the nine sub-scales of assessment literacy.

The results of pair-wise comparison tests (Table 6) indicated that among the 36 possible comparisons, only two showed significant differences between the sub-scales of assessment literacy.

A: The Iranian ESP computer sciences engineering teachers had significantly higher mean on "Using Assessment Results for Decision Making" ($M = 1.65$) than the sub-scale of "Recognizing Unethical Assessment Practices" ($M = 1$) ($MD = .65$, $p = 0.027$).

Table 5

Multivariate Tests; Sub-Scales of Assessment for Iranian ESP Teachers of Computer Sciences Engineering

		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
ESP	Pillai's Trace	.426	2.93	8	32	.013	.426
	Wilks' Lambda	.574	2.93	8	32	.013	.426
	Hotelling's Trace	.741	2.93	8	32	.013	.426
	Roy's Largest Root	.741	2.93	8	32	.013	.426

B: The Iranian ESP computer sciences engineering teachers had significantly higher mean on "Developing Assessment Methods" (M = 1.57) than the sub-scale of "Recognizing Unethical Assessment Practices" (M = 1) (MD = .575, p = 0.009).

Table 6

Pair-wise Comparisons; Sub-Scales of Assessment Literacy for Iranian ESP Teachers of Computer Sciences Engineering

(I) assessment	(J) assessment	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval for Difference	
					Lower Bound	Upper Bound
Choose AA	DevelopA	-0.07	0.197	1	-0.755	0.605
	AdminSI	-0.05	0.193	1	-0.713	0.613
	DevelopVGP	0.07	0.166	1	-0.495	0.645
	UsingARDM	-0.15	0.222	1	-0.915	0.615
	RecognizingUAP	0.50	0.16	0.121	-0.051	1.051
	KeepARA	-0.07	0.162	1	-0.632	0.482
	EnsureQM	0.05	0.199	1	-0.636	0.736
	CommunicateAR	-0.02	0.177	1	-0.635	0.585
Develop A	ChooseAA	0.75	0.197	1	-0.605	0.755
	AdminSI	0.02	0.158	1	-0.519	0.569
	DevelopVGP	0.15	0.188	1	-0.497	0.797
	UsingARDM	-0.07	0.158	1	-0.618	0.468
	RecognizingUAP	.575*	0.143	0.009	0.084	1.066
	KeepARA	0.00	0.156	1	-0.537	0.537
	EnsureQM	0.125	0.144	1	-0.371	0.621
	CommunicateAR	0.05	0.168	1	-0.528	0.628
Admin SI	ChooseAA	0.05	0.193	1	-0.613	0.713
	DevelopA	-0.02	0.158	1	-0.569	0.519
	DevelopVGP	0.12	0.183	1	-0.506	0.756
	UsingARDM	-0.10	0.202	1	-0.795	0.595
	RecognizingUAP	0.55	0.186	0.188	-0.090	1.190
	KeepARA	-0.02	0.166	1	-0.596	0.546
	EnsureQM	0.10	0.192	1	-0.562	0.762
	CommunicateAR	0.02	0.181	1	-0.597	0.647
Develop VGP	ChooseAA	-0.07	0.166	1	-0.645	0.495
	DevelopA	-0.15	0.150	1	-0.666	0.366

	AdminSI	-0.12	0.183	1	-0.756	0.506
	UsingARDM	-0.22	0.228	1	-1.009	0.559
	RecognizingUAP	0.42	0.202	1	-0.271	1.121
	KeepARA	-0.15	0.188	1	-0.797	0.497
	EnsureQM	-0.02	0.225	1	-0.800	0.75
	CommunicateAR	-0.10	0.155	1	-0.634	0.434
Using ARDM	ChooseAA	0.15	0.222	1	-0.615	0.915
	DevelopA	0.07	0.166	1	-0.495	0.645
	AdminSI	0.10	0.202	1	-0.595	0.795
	DevelopVGP	0.22	0.228	1	-0.559	1.009
	RecognizingUAP	0.65*	0.177	0.027	0.059	1.261
	KeepARA	0.07	0.158	1	-0.468	0.618
	EnsureQM	0.20	0.172	1	-0.594	0.794
	CommunicateAR	0.12	0.180	1	-0.494	0.744
Recognizing UAP	ChooseAA	-0.50	0.160	0.121	-1.051	0.051
	DevelopA	-.57*	0.143	0.009	-1.066	-0.084
	AdminSI	-0.55	0.186	0.188	-1.190	0.090
	DevelopVGP	-0.42	0.202	1	-1.121	0.271
	UsingARDM	-.65*	0.177	0.027	-1.261	-0.059
	KeepARA	-0.57	0.168	0.051	-1.152	0.002
	EnsureQM	-0.45	0.147	0.147	-0.958	0.058
	CommunicateAR	-0.52	0.168	0.12	-1.103	0.053
Kepp AR	ChooseAA	0.07	0.162	1	-0.482	0.632
	DevelopA	0.00	0.156	1	-0.537	0.537
	AdminSI	0.02	0.166	1	-0.546	0.596
	DevelopVGP	0.15	0.150	1	-0.366	0.666
	UsingARDM	-0.07	0.166	1	-0.645	0.495
	RecognizingUAP	0.57	0.168	0.051	-0.002	1.152
	EnsureQM	0.12	0.157	1	-0.415	0.665
	CommunicateAR	0.05	0.147	1	-0.458	0.558
Ensure QM	ChooseAA	-0.05	0.199	1	-0.736	0.636
	DevelopA	-0.12	0.157	1	-0.665	0.415
	AdminSI	-0.10	0.192	1	-0.762	0.562
	DevelopVGP	0.02	0.225	1	-0.750	0.800
	UsingARDM	-0.20	0.172	1	-0.794	0.594
	RecognizingUAP	0.45	0.147	0.17	-0.058	0.958
	KeepARA	-0.12	0.144	1	-0.621	0.371
	CommunicateAR	-0.07	0.166	1	-0.645	0.495
Communicate AR	ChooseAA	0.02	0.177	1	-0.585	0.635
	DevelopA	-0.05	0.147	1	-0.558	0.458
	AdminSI	-0.02	0.181	1	-0.647	0.597
	DevelopVGP	0.10	0.155	1	-0.434	0.634
	UsingARDM	-0.12	0.180	1	-0.744	0.494
	RecognizingUAP	0.52	0.168	0.12	-0.053	1.103
	KeepARA	-0.05	0.168	1	-0.628	0.528
	EnsureQM	0.07	0.166	1	-0.495	0.645

* The mean difference is significant at the .05 level.

Note: Each comparison is reported twice in Table 6. That is why half of the comparisons are negative.

6. Discussion

As noted earlier, teachers' assessment literacy is an important factor that has significant effect on students' learning and achievement. Moreover, knowing about the assessment literacy level of teachers for finding their strengths and weaknesses in using classroom assessment activities is of high importance.

Teaching English in higher education levels should be in accordance with the specific purpose of its major. Sometimes the decision makers believe that EFL teachers do not have the necessary knowledge about the social English of the field. On the other hand, specialist FSP university teachers should have the ability and quality of EFL teachers in addition to the knowledge of ESP (Hutchinson & Waters, 1987; as cited in Davoudi-Mobarakeh et al., 2014). Thus, to address the first research question, the specialist ESP teachers' means on the nine sub-scales against the criterion means of 2.50 was compared through parametric test of one-sample t-test. Since all p-values =.000 that are lower than the lowest corrected alpha level (0.05/8 = 0.006), and since the total mean of all sub-scales of specialist ESP teachers (M = 1.16) was significantly lower than the criterion mean of 2.50, it could be concluded that Iranian specialist ESP university teachers do not have acceptable level of assessment literacy.

As the results revealed, Iranian specialist ESP university teachers may be experts in their fields but they are not language experts and they do not have enough English skills. The results are also in line with Stiggins (1988) who believes that teachers are neither trained nor prepared to perform their role in the assessment process. Likewise, it also correlates with Davidheiser's (2013) opinion who states "few teachers are prepared to face the challenges of classroom assessment because they have not been given the opportunity to learn to do so" (p. 28)

In a nutshell, the results proved that Iranian ESP teachers of computer sciences engineering have the expected woefully low levels of assessment literacy and on average; they were not successful in overall assessment. It is the time for the Iranian educational system to be aware that teachers are not well-prepared to successfully assess their students learning and achievement. Actually, Iranian educational system pays limited attention to teachers needs in the preparation of language teachers and also pays limited attention to either the particular discourse of academic disciplines or to the practical concerns of needs analysis, text adaptation, curriculum development, or collaborative teaching in most language teacher training programs (Crandall, 2000).

To answer the second research question, a Multivariate Analysis of Variances (MANOVA) was used to compare the Iranian specialist ESP university teachers' means on the nine sub-scales of assessment literacy. Since the P-value =0.516 and $F(8, 32) = 0.917$, it can be concluded that there were no significant differences between the Iranian specialist ESP university teachers' means on their nine sub-scales of assessment literacy.

The result has helped to identify areas of weaknesses and strengths and showed that out of the nine sub-scales, the highest overall performance for all sub-scales of teachers' assessment was found for scale 8-Ensuring Quality Management (M =1.375) and the lowest performances were found for scale 2-Developing Appropriate Assessment Method and scale 3-Administrating, Scoring and Interpreting Assessment Results (M = 1.025). Therefore, as the results revealed, specialist ESP teachers had poor performance on all sub-scales of assessment literacy and they are not more efficient and qualified for assessing students' achievement which may demotivate learners. The main reason for such a gap is

insufficient knowledge of assessment or unfamiliarity of specialist ESP university teacher with implementing assessment in ESP classes.

These findings correlate with those of Estaji et al.'s (2024) results who found that teachers' assessment literacy is an important factor that has significant effect on students' learning and achievement. Moreover, knowing about the assessment literacy level of teachers for finding their strengths and weaknesses in classroom assessment activities is of high importance (Chang, 2024).

The present study was an attempt to examine the assessment literacy level of Iranian ESP computer sciences engineering teachers. Many of the results of this study were in line with earlier studies (Plake & Impara, 1993) that used the instrument which focused on the assessment literacy of teachers. Comparing the overall level of assessment literacy of Iranian ESP computer sciences engineering teachers against the criterion mean of 2.50, indicated that they do not have acceptable level of assessment literacy. The low level of assessment literacy among computer sciences engineering teachers may be due to "traditional teacher preparation courses in classroom assessment are not well matched with what teachers need to know for classroom practice" (Schafer, 1993; as cited in Davidheiser, 2013, p. 85). Furthermore, they did not receive sufficient training in their undergraduate preparation program.

In the present study, the highest mean performance for Iranian ESP teachers of computer sciences engineering was on the fifth scale, i.e., Using Assessment Results for Decision Making; the lowest was on the sixth scale, i.e., Recognizing Unethical Assessment Practices. Furthermore, comparisons between assessment literacy levels of Iranian ESP teachers of computer sciences engineering on the nine sub-scales' means revealed that there are significant differences on six of nine sub-scales, as well as on the total means of assessment literacy for this group. This may be due to Iranian ESP teachers of computer sciences engineering being unaware of the influential role of assessment in students learning achievement, the lack of courses in measurement and assessment in their education coursework, and inappropriateness of their assessment training.

7. Conclusions and Implications

The current investigation tried to assess the level of assessment literacy among Iranian ESP university instructors. A number of findings from this research were consistent with prior studies (Plake & Impara, 1993) utilizing an assessment literacy tool for educators. The comparison of Iranian ESP university teachers' overall assessment literacy level with the standard mean of 2.50 revealed a deficiency in their competence in this area. The inadequate assessment literacy observed in these teachers could be attributed to the discrepancy between traditional teacher training programs in assessment and the actual requirements for effective classroom assessment practices (Davidheiser, 2013, p. 85). Additionally, their undergraduate preparation did not adequately equip them with the necessary skills.

In the current investigation, the top average achievement among Iranian ESP university instructors was observed in scale 3, which pertains to Administering, Scoring, and Interpreting Assessment Results; conversely, the lowest performance was recorded in scale 8, focusing on Ensuring Quality Management. Similarly, the highest mean score for Iranian specialist ESP university teachers in the field was in scale 8, specifically related to Ensuring Quality Management; in contrast, the lowest scores were seen in scale 3, concerning Administering, Scoring, and Interpreting Assessment Results, and scale 2, addressing Developing Assessment Method.

Moreover, the comparison of assessment literacy levels among Iranian ESP university instructors across the nine sub-scales' averages indicated significant variances on six out of the nine sub-scales, in addition to the overall assessment literacy averages for this cohort. This divergence could potentially stem from a lack of awareness among Iranian ESP university teachers regarding the pivotal role of assessment in students' academic performance, the absence of modules focusing on measurement and assessment in their educational curriculum, and the inadequacy of their assessment preparation.

In a nutshell, the results proved that Iranian ESP computer sciences engineering teachers have the expected woefully low assessment literacy levels;and they generally were not successful assessment. It is the time for the Iranian educational system to be aware that teachers are not well-prepared to successfully assess their students learning and achievement. The results of this study will help the teachers to be aware of their assessment literacy level and to develop their classroom assessment literacy. Developing the assessment literacy level enables teachers to discover and evaluate their knowledge, understanding and values which shape their conceptualization, interpretations, judgments and decisions in assessing students' learning. By doing so, teachers should be able to know about the learning aspects to be assessed and for the reasons and the way that they can use the assessment results to increase their instruction and student's achievement and learning. The results of this study also encourage teachers to think deeply about the selection of assessment method, assessment purposes, grading procedures and the main stages of the assessment process and consequently, improving their level of understanding and learning in assessment and promoting their classroom assessment literacy.

Acknowledgment

We all thank the participants for their time, energy, and cooperation.

Authors' Contributions

All authors have conducted the study, collected data, analyzed and interpreted the data, and written up the manuscript.

Funding

The study did not receive any funding.

Competing Interests

The authors declare that there is no conflict of interest.

References

Alderson, J. C. (2000). *Assessing reading*. Cambridge University Press.

Bayat, Kh. & Rezaei, A. (2015). Importance of teachers' assessment literacy. *International Journal of English Language Education*, 3(1), 159–146. https://doi.org/10.5296/ijele.v3i1.6887

Brown, H. D. (2004). *Language assessment: Principles and classroom practices.* Pearson Education.

Crandall, J. (2000). Language teacher education. *Annual Review of Applied* Linguistics, 20(1), 34–55.

Çakır, N. C., & Genç, Z. S. (2022). A comparative analysis of teachers' beliefs about the assessment of 4th grade-EFL students in Turkey, Italy and Finland. *European Journal of Foreign Language Teaching, 5*(6). http://dx.doi.org/10.46827/ejfl.v5i6.4104

Chang, D. Y. S., Lin, M. H., & Lee, J. Y. (2024). Exploring language assessment literacy and needs of English teachers at senior high school level. *Asia Pacific Journal of Education*, 1–19. https://doi.org/10.1080/02188791.2024.2313486

Davidheiser, S. A. (2013). *Identifying areas for high-school teacher development: A study of assessment literacy in the Central Bucks school district* [A published master's thesis]. School of education at Drexel University.

Davies, A. (2008). Textbook trends in teaching language testing. *Language Testing, 25*(3), 327–347. https://doi.org/10.1177/0265532208090156

Davoudi-Mobarakeh, S., Eslami-Rasekh, A., & Barati, H. (2014). Observation and feedback of content specialists versus general English teachers: Suggestions to make optimal English for specific purposes courses. *Journal of English Language Teaching and Learning, 6*(13), 15–41.

Elshawa, N. R., Heng, C. S., Abdullah, A. N., & Rashid, S. M. (2016). Teachers' assessment literacy and washback effect of assessment. *International Journal of Applied Linguistics and English Literature, 5*(4), 135–141. http://dx.doi.org/10.7575/aiac.ijalel.v.5n.4p.135

Estaji, M., Brown, G. T., & Banitalebi, Z. (2024). The key competencies and components of teacher assessment literacy in digital environments: A scoping review. *Teaching and Teacher Education, 141*(1), 104497. https://doi.org/10.1016/j.tate.2024.104497

Förster, C. E. (2023). Teacher assessment literacy. in: förster, c.e. (eds) the power of assessment in the classroom. Springer texts in education. *Springer, Cham*. https://doi.org/10.1007/978-3-031-45838-5_1

Fulcher, G. (2012). Assessment literacy for the language classroom. *Language Assessment Quarterly, 9*(2), 113–132. https://doi.org/10.1080/15434303.2011.642041

Giraldo, F. (2018). Language assessment literacy: Implications for language teachers. *Profile Issues in Teachers Professional Development, 20*(1), 179-195. http://dx.doi.org/10.15446/profile.v20n1.62089

Gkogkou, E., & Kofou, I. (2021). A toolkit for the investigation of Greek EFL teachers' assessment literacy. *Languages, 6*(4), 188. https://doi.org/10.3390/languages6040188

Harding, L., & Kremmel, B. (2016). Teacher assessment literacy and professional development. In D. Tsagari & J. Banerjee, *Handbook of second language assessment* (pp.413–428). https://doi.org/10.1515/9781614513827-027

Heaton, J. B. (2011). *Writing English language tests* (new edition). Longman Group.

Howell, C. A. (2013). *Development and analysis of a measurement scale for teacher assessment literacy* [MA Thesis]. Faculty of the Department of School Psychology East Carolina University.

Jan-nesar, M. Q., Khodabakhshzadeh, H., & Motallebzadeh, K. (2020). Assessment literacy of Iranian EFL Teachers: A Review of Recent Studies. *Journal of Asia TEFL, 17*(2), 689–705. http://dx.doi.org/10.18823/asiatefl.2020.17.2.27.689

McNamara, T. (2000). *Language testing*. Oxford University Press.

Mertler, C. A. (2003). Pre-service versus in-service teachers' assessment literacy: Does classroom experience make a difference? Paper presented at *the annual meeting of the Mid-Western Educational Research Association, Columbus, Ohio.*

Plake, B. S., & Impara, J. C. (1997). Teacher assessment literacy: What do teachers know about assessment? In G. D. Phye (Ed.), *Handbook of classroom assessment* (pp. 53–68). Academic Press.

Popham, W. J. (2009). Assessment literacy for teachers: Faddish or fundamental? *Theory into Practice, 48*(1), 4–11. https://doi.org/10.1080/00405840802577536

Rezai, A. (2024). The role of teacher assessment literacy in job stress and job burnout in EFL contexts: a mixed-methods investigation. *Asian-Pacific Journal of Second and Foreign Language Education, 9*(1), 3. https://doi.org/10.1186/s40862-023-00225-1

Scarino, A. (2013). Language assessment literacy as self-awareness: Understanding the role of interpretation in assessment and teacher learning. *Language Testing, 30*(3), 309–327. https://doi.org/10.1177/0265532213480128

Shohamy, E. (2001). *The power of tests: A critical perspective on the use of language tests*. Longman.

Stiggins, R. (1991). Assessment literacy. *Phi Delta Kappa, 72*(7), 534-559. https://doi.org/10.36348/sjhss.2021.v06i07.002

Stiggins, R. J. (1995). Assessment literacy for the 21st century. *Phi Delta Kappan, 77*(3), 238–245.

Stiggins, R. (1999). Are you assessment literate? *The High School Journal, 6*(5), 20–23.

Stiggins, R. (2000). Learning teams for assessment literacy. *Orbit, 30*(4), 5–7.

Stiggins, R. (2006). Assessment for learning: A key to motivation and achievement. *Edge: The Latest Information for the Education Practitioner, 2*(2), 1–19.

Tao, N. (2014). Development and validation of classroom assessment literacy scales: *English as a Foreign Language (EFL) instructors in a cambodian higher education setting* [Doctoral dissertation]. Victoria University.

Ukrayinska, O. (2024). Synergies in developing pre-service teachers' language assessment literacy in Ukrainian Universities. *Education Sciences*, *14*(3), 223. https://doi.org/10.3390/educsci14030223

Webb, N. L. (2002). Assessment literacy in a standards-based urban education setting. *Wisconsin Center for Education Research, 4*(1), 1–19.

Appendix

Classroom Assessment Knowledge Test

DIRECTIONS

The following items are examining your knowledge in the educational assessment of students. Please read each scenario followed by each item carefully and answer each of the items by circling the response you think is the best one. Even if you are not sure of your choice, circle the response you believe to be the best. Do not leave any items unanswered.

Scenario # 1

Mr. Ahmady a first year English writing lecturer, is aware of the fact that his students will be taking a semester examination at the end of the course.

1. Mr. Ahmady wants to assess his students‟ critical thinking abilities at the end of the unit to determine if any reinstruction will be necessary prior to the exam. Which of the following methods would be the most appropriate choice?

A. multiple-choice items

B. matching items

C. gap-filling items

D. essay writing

2. In order to grade his students' writing accurately and consistently, Mr.Ahmady would be well advised to

A. identify criteria from the unit objectives and create a marking criteria.

B. develop a marking criteria after getting a feel for what students can do.

C. consider student performance on similar types of tests.

D. consult with experienced colleagues about a marking criteria that has been used in the past.

3. Mr. Ahmady wants to evaluate his students' understanding of specific aspects of their responses. Which of the following would best facilitate him scoring of these responses?

A. an objective answer key

B. a holistic scoring

C. a checklist

D. an analytic scoring

4. At the end of each class period, Mr. Ahmady asks his students several questions to get an impression of their understanding. In this example, the primary purpose for conducting formative assessment is to

A. determine the final grades for students.

B. determine content for the final examination.

C. identify individual learning needs to plan classroom instruction.

D. evaluate curriculum appropriateness.

5. Which grading practice being considered by Mr. Ahmady would result in grades that would most reflect his students‟ learning achievement against the learning outcomes?

A. grades based on the students‟ performances on a range of assessments

B. grades based on the amount of time and effort the student spent on the assessments

C. grades based on how the student has performed in comparison to his/her classmates

D. grades based upon the personal expectations of Mr. Ahmady

6. Mr. Ahmady is planning to keep assessment records as a part of his assessment and reporting process. Which of the following is the least important assessment information to be recorded?

A. statistical data including marks, student welfare and biographical information.

B. anecdotal data comprising critical incidents or reflections of both Mr. Ahmady and his students.

C. all copies of his students‟ assessment work.

D. a representative sample of each student work.

7. In a routine conference with his students, Mr. Ahmady is asked to explain the basis for assigning his course grade. Mr. Ahmady should

A. explain that the grading system was imposed by the school administrators.

B. refer to the information that he presented to his students at the beginning of the course on the assessment process.

C. re-explain the students the way in which the grade was determined and show them samples of their work.

D. indicate that the grading system is imposed by the Ministry of Education.

8. Mr. Ahmady as worried that his students would not perform well on the semester examination. He did all of the following to help increase his students‟ scores. Which was unethical?

A. He instructed his students in strategies for taking tests.

B. He planned his instruction so that it focused on concepts and skills to be covered on the test.

C. He allowed his students to bring in their course books/materials to refer to during the test.

D. He allowed students to practice with a small number of items from the actual test.

9. To ensure the validity and reliability of his classroom assessment procedure, it is advised that Mr. Ahmady should gather together with his colleagues to discuss all of the following except

A. marking criteria.

B. Students’ pieces of work.

C. teaching techniques.

D. assessment activities.

Scenario # 2

Ms. Nasiry is a year two English lecturer. She has just finished teaching a unit on the Industrial Revolution and wishes to measure her students’ understanding of this particular unit using a multiple-choice test.

10. Based on her goal, which of the following assessment strategies would be the most appropriate choice?

A. She should use the test items included in the teacher’s manual from the textbook she uses.

B. She should design test items which are consistent with the content and skill specified in the course learning outcomes.

C. She should use available test items from internet that cover Industrial Revolution.

D. She should design test items which cover the factual information she taught.

11. In constructing her multiple-choice test items, Ms. Nasiry all of the following guidelines except

A. ensure that the correct response is unequivocally the best.

B. ensure that the responses to a given item are in different literary forms.

C. ensure the stem and any response, taken together, read grammatically.

D. make all distracters plausible and attractive to the ignorant test-taker.

12. Ms. Nasiry decides to score the tests using a 100% correct scale. Generally speaking, what is the proper interpretation of a student score of 85 on this scale?

A. The student answered 85% of the items on the test correctly.

B. The student knows 85% of the content covered by this instructional unit.

C. The student scored higher than 85% of other students who took this test.

D. The student scored lower than 85% of other students who took this test.

13. Some of Ms. Nasiry's students do not score well on the multiple-choice test. She decides that the next time she teaches this unit; she will begin by administering a pretest to check for students‟ prerequisite knowledge. She will then adjust her instruction based on the pretest results. What type of information is Ms. Nasiry using?

A. norm-referenced information (describes each student's performance relative to the other students in a group such as percentile ranks)

B. criterion-referenced information (describes each student's performance in terms of status in specific learning outcomes)

C. both norm- and criterion-referenced information

D. neither norm- nor criterion-referenced information

14. The Industrial Revolution test is the only student work that Ms. Nasiry grades for the current grading period. Therefore, grades are assigned only on the basis of the test. Which of the following is not a criticism of this practice?

A. The test, and therefore the grades, reflects too narrow a curriculum focus.

B. These grades, since based on test alone, are probably biased against some minority students.

C. Tests administered under supervised conditions are more reliable than those assessments undertaken in less standardized conditions (e.g., homework)

D. Decisions like grades should be based on more than one piece of information

15. Ms. Nasiry fully understands that her classroom assessment records serve all of the following purposes except

A. provide information regarding assessment methods development.

B. provide diagnostic information to show the strengths and weaknesses of student performance.

C. show the extent of student progress.

D. provide information to assist administrative decision makers.

16. During an individual conference, one student in Ms. Nasiry class wants to know what it means that he scored in the 80th percentile in a multiple-choice test. Which of the following provides the best explanation of this student's score?

A. He got 80 % of the items on the test correct.

B. He is likely to earn a grade of "B" in his class.

C. He is demonstrating above grade level performance.

D. He scored the same or better than 80 % of his classmates.

17. Based on their grades from last semester, Ms. Nasiry believes that some of her low-scoring students are brighter than their test scores indicate. Based on this knowledge, she decides to add some points to their test scores, thus raising their grades. Which of Ms. Nasiry 's action was unethical?

A. examining her student's previous academic performance

B. adjusting grades in her course

C. using previous grades to adjust current grades

D.adjusting some students‟ grades and not others‟

18. To enhance the quality of a new developed multiple-choice test, Ms. Nasiry should do all of the following except

A. pilot the test items with a small number of her past students to see how well each item performs.

B. make all necessary changes to the test items based on the information received during her pilot.

C. have all of her current students undertake the test twice and make a comparison of their scores.

D. panel the test items through consultation with her colleagues who have assessment experience.

19. Scenario # 3

Mr. Farhady is a senior English lecturer in the IUST University. Experienced in issues of classroom assessment, Mr. Farhady is often asked to respond to the questions concerning best practices for evaluating student learning.

19. Ms. Nasiry, an English lecturer, asks what type of assessment is best to determine how well her students are able to apply what they have learned in class to a situation encountered in their everyday lives. The type of assessment that would best answer her question is called

A. diagnostic assessment.

B. performance assessment.

C. formative assessment.

D. authentic assessment.

20. Ms. Nasiry is constructing essay questions for a test to measure her students‟ critical thinking skills. She consults with Mr. Farhady to see what concerns she would be aware of when constructing the questions. Which statement is not an appropriate recommendation when writing essay questions?

A. consider the relevance of the questions for a particular group of her students

B. avoid determining the amount of freedom of writing responses that will be accepted

C. indicate the time limits for the writing responses

D. be clear about the skills require to be demonstrated

21.Roya, a student in Mr. Farhady's class, scored 78 marks on a reading test which has a mean of 80 and a standard deviation of 4. She scored 60 marks on the writing test which had a mean of 50 and a standard deviation of 3. Based on the above information, in comparison to her peers, which statement provides the most accurate interpretation?

A. Roya is better in reading than in writing.

B. Roya is better in writing than in reading.

C. Roya is below average in both subjects.

D. Roya is close to average in both subjects.

22. After teaching four units from his course book, Mr. Farhady gives his students a test to measure their learning achievement. In this example, the primary purpose for conducting summative assessment is to

A. identify individual learning needs to plan classroom instruction.

B. motivate students to learn.

C. evaluate curriculum appropriateness.

D. determine the final grades for students.

23. Throughout instruction, Mr. Ahmady assesses how well his students are grasping the material. These assessments range from giving short quizzes, mid-term tests, written assignments to administering a semester examination. In order to improve the validity of this grading procedure, what advice should Mr. Farhady give to Mr. Ahmady?

A. consider students‟ class participation and their attendance before assigning a final grade.

B. consider students‟ performance in other subjects before assigning a final grade.

C. weight assessments according to their relative importance.

D. take into consideration each student's effort when calculating grades.

24. Ms. Azad consults with Mr. Farhady for advice to effectively use her observations in recording her students‟ activities in the classroom. Which statement is not an appropriate recommendation when observing her students‟ behaviors?

A. make a record of the incident as soon after the observation as possible

B. maintain separate records of the factual description of the incident and her interpretation of the event

C. observe as many incidents in one long observation as possible

D. record both positive and negative behavioral incidents

25. Boshra is a student in Mr. Farhady's class. He receives a raw score of 12 items answered correctly out of a possible 15 on the vocabulary section of a test. This raw score equates to a percentile rank of 45. He is confused about how he could answer so many items correctly, but receive such a low percentile rank. He approaches Mr. Farhady for a possible explanation. Which of the following is the appropriate explanation to offer to Bora?

A. "I don't know…there must be something wrong with the way the test is scored."

B. "Although he answered 12 correctly, numerous students answered more than12 correctly."

C. "Raw scores are purely criterion-referenced and percentile ranks are merely one form of norm-referenced scoring."

D. "Raw scores are purely norm-referenced and percentile ranks are merely one form of criterion-referenced scoring."

26. Prior to the semester examination, Mr. Farhady reveals some information to his students. Which of Mr. Farhady's action was unethical?

A. inform his students the exam contents to be covered.

B. inform his students the exam methods to be used.

C. show the actual exam paper to a small group of his low-achieving students.

D. tell his students the exam duration.

27. To achieve quality management of classroom assessments, Mr. Frahady advises his colleagues to be involved in all of the following except

A. quality assurance (concerning with quality of assessment by emphasizing the assessment process).

B. quality teaching (dealing with the effectiveness of teaching in helping students undertake assessments successfully).

C. quality control (dealing with monitoring and, where necessary making adjustment to assessor judgments before results are finalized).

D. quality review (focusing on the review of the assessment results and processes in order to make recommendations for future improvement).