ASSESSMENT FOR LEARNING AT SAUDI UNIVERSITIES: AN ANALYTICAL STUDY OF ACTUAL PRACTICES

Rommel M. A. AlAli
King Faisal University, Al-Ahsa, Saudi Arabia

ABSTRACT

Assessment for learning is a part of learning and teaching processes, diagnosing learners' needs, and providing them with feedback to improve their performance. This study is aimed to investigate and analyze the practices of assessment for learning among the faculty members at Saudi Universities. This study uses a quantitative survey approach. Psychometric properties of assessment for learning instrument using Rasch Model Analysis were verified on a sample of (255) faculty members from Saudi Universities. The findings showed that the practice degree of assessment for learning among the faculty members was medium. In addition, there were no statistically significant differences in the practice degree of faculty members of assessment for learning according to gender, faculty, and teaching experience. While there were statistically significant differences in all dimensions of assessment for learning based on academic rank. Finally, the study recommended that conducting training programs on strategies of assessment for learning and holding workshops to exchange experiences between all faculty members.

Keywords – Rasch Model Analysis, demographic factors, Psychometric properties

Introduction

Considering the scientific and technological development and progress, it is essential to change and renew the educational process. Recently, the focus has become on constructivism visions, which rejects to divide the curriculum into separate skills. It emphasizes that the mind creates knowledge, and the teacher is a supporter in building that knowledge (Rizk, 2014; Alsbeeh, 2017; Flórez & Sammons, 2013; Alkhayal, 2019). Since the assessment is one of the components of the educational process, hence it has an important role in the learning and teaching process because it considers the main source to make decisions that are related to learning difficulties and their diagnosis, it is also encouraging and enhance student' learning. Furthermore, it considers an effective tool to judge on progress of students, curriculum, programs, and educational policies. There is a clear and strong relationship between assessment and teacher's teaching methods and student learning styles. Educational assessment plays an important role in determining the level of learning, and in the appropriate education, methods to achieve its goals (Ryan, 2015; Alshamekh, 2018; Akib and Ghafar, 2019; Taras, 2010).

Assessment is a systematic process for collection, analyzing, and using information from the outcomes that were measured in an organized and continuous method to improve student learning (Akib and Ghafar, 2019; Darandari, 2017). Considering the call for developing the educational process and adopting comprehensive quality, the evaluation process should be reformed because of its importance within the educational system. The ’fixed educational concepts for teachers must be modified. Assessment is a tool that the teacher uses to judge students, it is not considered as a way to engage them in a constructive assessment environment, however, it is to develop a positive trend towards the evaluation process. The assessment is necessary for educational institutions, but the need for it is more in universities because its application successfully leads to achieving the desired goals, such as raising the academic level, developing creativity, and achieving communication between the aspects of the educational process (Azizi, 2018; Alkhayal, 2019).

Assessment for learning (AFL) is often used to describe constructive assessment strategies. It is a formative and diagnostic assessment integrated into the teaching and learning process to continuously modify strategies of this process. It focuses on developing the quality of learning and improving performances (Alkhayal, 2019; Abdulkareem & Omer, 2019; Azizi, 2018; Darandari, 2017; Ryan, 2015). Council of Chief State School Officers (2018) defined AFL as a planned and continuous process, used by all teachers and students during the learning and teaching process, to extract and elicit evidence about student learning, use it to improve student understanding of targeted learning outcomes, and support students to become self-learners. Therefore, AFL is interested in employing various methods of assessment and using its data by students, teachers, and parents in the development of quality of learning. Students learn better which makes them the focus of the evaluation process and it helps them to practice and feel them be able to control and achieve success by continuing to work on tasks and activities (Stiggins,

2005). AFL is an effective way to raise student achievement. Students' performance will improve significantly if they understand the goal of their learning. The characteristics of assessment for learning represent the general traits of a good teacher. This requires the teacher to obtain more details about the progress students are making in achieving the curriculum goals to think in different ways to help them (Jawawi et al., 2017; Akib & Ghafar, 2015; Flórez & Sammons, 2013; Haroldson, 2012). AFL helps the teacher to follow the learner's growth in the cognitive, emotional, and psychomotor fields, provide him with feedback on students' mistakes , and provides him with the appropriate data on their progress and level of achievement (Azizi, 2018).

AFL has several basic principles; it covers all aspects of achievement for all students, develop their ability to peer and self – assessment, helps them to know how to improve students', supports understanding of learning outcomes and assessment criteria, builds and supports motivation, meaningful, targeted, sensitive, and takes into account the emotional aspects, essential for professional development, essential for classroom practices, focuses on how the student learns, and it is considered part of effective learning and teaching planning (Darandari, 2017; Jones & Saville, 2016).

AFL is used not only to confirm the learning occurring but to raise the level of learning. It is multi-dimensional, constructional, integrated into the curriculum, real and flexible used in an early stage of learning to diagnose the needs of students. This type provides information about student learning, and the effectiveness of the learning strategies they use (Darandari, 2017; Stiggins & Chappuis, 2013; Arends & Kilcher, 2015).

AFL determines the learners' levels and status from their learning processes, enhances their learning, and makes them aware of their strengths and weaknesses, to improve the learning and teaching process. (AFL) provides an environment rich with feedback, which is done through the teacher's dialogue with his students, and the interaction of students with each other. This environment provides students with opportunities to apply their knowledge, skills, and understanding to learning content. It also provides them with opportunities to improve and progress towards achieving their learning goals. Moreover, helping students to engage in their careers in the future (Ryan, 2015; Pang and Leung, 2011; McDowella et al., 2011).

AFL judges the quality of learning, to determine the next steps of action. It is designed to assess both students and teachers. It also uses clear and detailed descriptive feedback, focuses on improvement and comparison with students' previous performance, promotes students' success beliefs, and helps them get rid of fear and dread that are due to the traditional assessment practices. Furthermore, it helps students on follow-up and adjust learning methods, improves their level of mastery of educational content, increase their performance on all tests, positively affects students' perception of their learning, and enhances their motivation and aptitude to learn (Chappuis et al., 2011; Popham & Stiggins, 2007).

One of the strategies of assessment for learning is that students work in cooperative groups, where peers assess each other by sharing their ideas, suggestions, and decisions, and judging them after comparing them to the success criteria. After that, the misconceptions change directly between the group members due to the exchange of ideas and the diversity of experiences. Peer assessment provides an opportunity for every student to adjust and improve performance; it raises the quality and increases academic achievement (Keeley, 2015). Evidence about students' understanding process can be gathered through various tools, such as listening to student discussions, observing their performance, through written work, learning assignments, or tests of all kinds. The information gathered about student learning should be analyzed by the teacher or the student himself, then work out procedures to improve learning, thus it becomes an assessment for learning (Shepard, 2000).

The results of the multiple studies that were aimed to identify and investigate the assessment practices in Saudi Universities indicated that there is a variation in the degree of assessment practices (Otaibi, 2018; Alshamrani, 2017; Alsbeeh, 2017). There are also some misunderstandings of the main assessment concepts, there are problems in practicing them, and it is refraining from changing the assessment practice. In addition, some studies conducted training programs and suggested models of assessment for learning to improve the practice of applying and practicing its tools and strategies (Azizi, 2018; Abdulkareem & Omer, 2019; Abdulkareem & Alshaya, 2018). The constructivist view of learning focuses on the learner's role in learning and assessment processes and practices (Darandari, 2017 and 2014).

AFL is one of the most prominent recent trends in educational assessment, and the knowledge of assessment practices for learning enables faculty members to take advantage of it in their teaching behavior and enrich research, hence this study came to develop a tool that includes appropriate assessment practices for learning. The scale is used to measure the degree of the practices of assessment for learning by faculty members.

Literature Review

The study by Darandari (2017) discussed the characteristics and strategies of effective assessment for learning and ways to implement it in the classroom to enhance learning, in addition to developing policies and establishing effective systems for assessment for learning. The study by Gilles, Detroz, & Blais (2011) aimed to investigate classroom assessment practices for faculty members in higher education. The assessment practices of the participating universities from different countries are displayed on the online platform, and these practices are compared with each other. Zacharis (2010) conducted a study that examined the effect of motivation for assessment on student achievement. It also focused on innovative assessment for learning methods to improve learning. Several studies have shown weaknesses and deficiencies in evaluation practices in general and weak practices of teachers for formative assessment.

Therefore, it is necessary to implement the "assessment for learning" effectively to improve students learning (Albursan et al., 2015; Volante & Beckette, 2011; Burns, 2010; Jett, 2009).

Many studies were aimed to identify and investigate the assessment for learning practices of teachers in schools. The results of these studies indicated that there is a variation in the degree of assessment practices from low to high. It also showed a provision of all kinds of written tests (Otaibi, 2018; Alshamrani, 2017; Alsbeeh, 2017; Ryan, 2015; Albursan et al., 2015; Al-Bashir and Barham, 2013; Sharah and Zaza, 2013; Refaee et al., 2012; Zhang & Burry-Stock, 2003; Mcmillan et al., 2002;). To obtain positive and good results of assessment for learning, some studies have relied on Rasch model analysis as a tool of assessment for learning, due to its accuracy and effectiveness in developing items of tests and tools (Akib and Ghafar, 2019; Sumintono, 2018; Akib & Ghafar, 2015). Given the importance of assessment for learning in the learning and teaching process. Some studies conducted training programs and suggested models of assessment for learning to improve the level of achievement, and the practice of applying its tools and strategies (Alkhayal, 2019 ;Abdulkareem & Omer, 2019; Abdulkareem & Alshaya, 2018; Abdulkareem & Alshaya, 2016).

To provide a class environment based on assessment for learning, teachers should change their assessment practices. The most important practices are; focusing on learning by sharing learners to determine the learning objectives, providing effective questions that enhance learners' thinking skills, providing effective feedback that includes clear guidelines and directed at improving student learning, and peer and self – assessment which allow students to discuss their learning, and discuss the level of their awareness and mistakes with colleagues (Alkhayal, 2019; Akib and Ghafar, 2019; Alsbeeh, 2017; Darandari, 2017; Ryan, 2015; Akib & Ghafar, 2015; Erwin and Najib, 2015; Flórez & Sammons, 2013; and Gardner, 2009).

The teacher’s awareness of the level of assessment culture is low, misunderstanding of the main assessment concepts, there are problems in applying it, and refrain from changing the assessment practice, which leads to failure to achieve goals of assessment for learning (Volante & Beckett, 2011; Kanjee & Mthembu, 2015; Darandari and Murphy, 2013; Abdulkareem & Omer, 2015). There are difficulties in applying the formative evaluation because of lack of time, increase in the number of students, and intensity of the curriculum (Mariam, 2016). This research paper aims to analyze the practices of assessment for learning among the faculty members at Saudi Universities. Furthermore, the research questions in this paper are: Is the developed instrument valid and reliable to measure practices of assessment for learning to faculty members? What is the practice degree of the faculty member for assessment for learning? Finally, are there any statistically significant differences in the level of practice of faculty members for assessment for learning according to gender, academic rank, teaching experience, and faculty?

Rasch Model Analysis

Rasch model enables teachers to develop test items and substantial tools, by providing psychometrics analysis methods and providing information related to students' assessment for learning (Sumintono, 2018). Rasch developed a special model, to estimate the abilities of individuals through their responses on the test items (McCamey, 2014). Rasch model helps to predict the probability of the correct answer on a test based on estimate items two variables which are the difficulty of item and ability of the individual through joint continuity between them (De Battisti et al., 2004). Rasch model analysis improves the accuracy and quality of tests and surveys, as it also allows the creation of multiple forms of measuring instruments. When using survey data, it makes important corrections and clarifies the meaning of student and group metrics using survey items (Boone, 2016). The construction of any achievement test according to the Rasch model provides the advantages of accuracy, objectivity, and independence in the measurement. Rasch model is taken as a criterion for the structure of the responses, rather than a mere statistical description of the responses. Rasch model is used to reach the highest level of accuracy and objectivity in the measurement to achieve a more accurate relationship between measuring tool and underlying attribute of the individual (Nunnally, 1994). Rasch Model analysis is a powerful tool for evaluating constructs validity and reliability of the instrument (Mofreh et al., 2014).

Methodology

This study used a quantitative descriptive survey approach. The population is comprised of all faculty members at Saudi Universities. King Faisal University has been specifically chosen from Saudi Arabia during the academic year 2019/2020. There was a total of 2,012 faculty members at the University during this time frame. A sample of this study has been randomly selected of all colleges at King Faisal University which consists of (255) faculty members. To achieve the objectives of this study and answer the research questions, the scale of assessment for learning practices was developed. The scale is used to measure the degree of the practices of assessment for learning by faculty members. Dimensions of the scale were determined by reviewing the past studies. It consisted of four dimensions namely the sharing learning objectives (SL), effective classroom questions (EQ), providing effective feedback (EF), and peer and self–assessment (PS). Items of the scale were developed by reviewing the books, the internet, and previous studies.

Verifying the validity and reliability of the instrument

The validity and reliability of the instrument were verified. Nine experts working at the University of King Faisal examined the instrument items. Based on their professional opinions, five instrument items were omitted, and some items were modified and reformulated. Approximately 30 respondents examined the instrument. The data was analyzed and evaluated

according to the Rasch model using Winsteps software version 3.68.2. The validity of the instrument was measured using values of MNSQ for infit, it should lie between 0.4 and 1.5, item polarity analysis (PTMEA), this value of PTMEA should lie between 0.2 and 1, standardized fit statistic (Zstd) value should be range between -2,2. Calibration scaling analysis, and the dimensionality, where the raw variance explained by measures should be more than 40% and unexplained variance in 1st contrast less than 15. The reliability of the instrument was measured using person and item reliability (Mofreh et al., 2017; Boone, 2016; Erwin and Najib, 2015).

Table 1: Item Fit Analysis of assessment for learning for the faculty members

Items	measure	Model	Infit		Outfit		Pt-measure
		S.E	MNSQ	ZSTD	MNSQ	ZSTD	CORR
PS10	.24	0.25	1.48	1.9	1.46	1.8	.33
PS8	1.53	0.42	1.44	1.9	1.44	1.4	.26
PS4	1.49	0.26	2.43	1.8	1.41	1.9	.23
PS7	-.29	0.41	1.42	1.8	1.45	1.5	.26
PS9	.11	0.29	1.25	1.7	1.37	1.0	.25
PS3	-.85	0.38	1.39	1.7	1.34	1.9	.22
PS2	-.92	0.34	1.37	1.7	1.34	1.3	.27
PS1	.87	0.25	1.39	1.5	1.34	1.6	.40
PS6	.66	0.36	1.28	1.1	1.30	1.2	.43
PS5	.73	0.28	1.28	1.5	1.48	1.1	.43
SL2	-.14	0.37	.89	-.4	1.01	.1	.60
SL1	-.17	0.34	.90	-.3	1.04	.2	.61
SL4	.20	0.35	.90	-.3	.85	-.6	.63
SL10	-.40	0.33	.98	.0	1.08	.4	.64
SL5	-.35	0.31	.94	.1	.99	.1	.66
EF2	-.20	0.35	.83	-.9	.80	-.5	.66
SL6	-.22	0.29	.97	-1.1	.90	-.2	.67
EQ11	-.40	0.33	1.00	-1.2	.97	.0	.68
EF12	.02	0.38	.78	-1.0	.73	-1.1	.69
SL7	.28	0.36	.74	-1.4	.75	-1.0	.71
EF4	.04	0.35	.72	-1.4	.67	-1.3	.73
EQ7	-.04	0.32	.70	-1.3	.70	-.9	.73
EF6	-.29	0.41	.67	-.9	.60	-1.5	.73
EF3	-.06	0.37	.69	-1.6	.65	-1.5	.74
EQ3	-.62	0.32	.67	-1.6	.61	-1.4	.75
EF5	-.11	0.29	.75	-1.4	.72	-.9	.75
EF9	.04	0.35	.66	-1.3	.62	-1.5	.76
EQ5	-.40	0.33	.61	-1.2	.63	-1.4	.76
EQ4	-.40	0.33	.63	-1.5	.60	-1.5	.76
EF11	-.46	0.30	.68	-1.3	.64	-1.3	.76
EF8	-.20	0.35	.66	-1.4	.57	-1.4	.77
EQ2	.77	0.37	.64	-1.9	.63	-1.5	.78
EF14	-.34	0.32	.67	-1.3	.62	-1.5	.78
SL3	-.16	0.32	.64	-1.9	.61	-1.6	.79
EQ1	-.14	0.37	.61	-1.5	.57	-1.9	.79
EQ9	.15	0.32	.63	-2.1	.57	-1.4	.79
EF10	.12	0.35	.60	-1.7	.57	-1.9	.80
SL9	-.46	0.32	.63	-1.9	.58	-1.7	.80
SL8	.39	0.35	.58	-1.8	.55	-1.9	.80
EF13	.28	0.30	.61	-1.5	.56	-1.7	.81
EF1	-.16	0.34	.54	-1.3	.52	-1.9	.83
EF7	.15	0.36	.54	-1.5	.50	-1.4	.84
EQ8	-.06	0.30	.43	-1.3	.44	-1.3	.85
EQ10	.12	.32	.43	-1.3	.42	-1.2	.86
EQ6	-.14	0.29	.44	-1.5	.45	-1.4	.86

The values of MNSQ for infit ranged from 0.44 to 1.48, as for the point measure correlation (PTMEA) value, all AFL items showed a positive value and greater than 0.20, this indicates that all items are moving in parallel function to measure the dimensions formed. The ZTS values ranged from -2 to +2 as shown in Table 1 below. These values are appropriate and acceptable for construct validity according to the Rasch model.

The dimensionality analysis result of AFL for the faculty members was illustrated in Table 2 below. The raw variance explained by measured value was 44.6%, which is more than 40%, and unexplained variance in 1st contrast value was 10.2%, which is less than 15. Thus, dimensionality data results in the post that the AFL data fit the Rasch model.

Table 2: Item Dimensionality of assessment for learning for the faculty members

	Empirical			Modeled
Total raw variance in observations	74.4	100%		100%
Raw variance explained by measures	39.4	49.5%		43.1%
Raw variance explained by persons	17.0	22.9%		25.0%
Raw Variance explained by items	12.4	16.6%		18.1%
Raw unexplained variance (total)	45.0	60.5%	100.0%	56.9%
Unexplained var.in 1st contrast	7.6	10.2%	16.9%	13.9%
Unexplained var.in 2nd contrast	5.3	7.1%	11.7%	11.7%
Unexplained var.in 3rd contrast	4.3	5.8%	9.5%	9.8%
Unexplained variance in 4th contrast	3.5	4.6%	7.7%	8.5%
Unexplained variance in 5th contrast	3.2	4.2%	7.0%	7.6%

For grading scale calibration analysis of the AFL as shown in Table 3 below, the most frequent answer is the scale of participants ranking 4 which is 16 (53%). The second grading scale is scale 3 which is 11 (37%), and the last grading scale is scale 2 which is 3 (10%). The column of observed averages shows the pattern of faculty members' move from -.84 to +1.83. Based on the Rasch model this indicates a normal pattern.

Table 3: Calibration Scaling Analysis of assessment for learning for the faculty members

Category Lable	Score	Observed Count %	Observed Average	Infit MNSQ	Outfit MNSQ	Structure Calibration	Category Measure
2	2	3 10	-.84	.72	.64	None	(-2.40)
3	3	11 37	.78	.88	1.29	-1.04	-.17
4	4	16 53	1.83	1.06	1.07	1.04	2.05

The person reliability is 0.94, which is greater than 0.5. Furthermore, the person separation is 4.12, which is greater than 2 as shown in Table 4 below. Based on the Rasch model these reliability values indicate that the instrument has a good degree of reliability. The value of the item's reliability is 0.69, which is greater than 0.5. The values of item separation are 2.93, which

is greater than 2 as shown in Table 5 below. Based on the Rasch model these reliability values indicate that the instrument has a good degree of reliability.

Table 4: Person Separation and Reliability of assessment for learning for the faculty members scale

	Raw Score	Count	Measure	Error	Infit		Outfit	
					MNSQ	ZSTD	MNSQ	ZSTD
Mean	143.6	45.0	1.17	.28	1.10	0.3	1.14	.2
S.D.	19.5	.0	1.34	.05	0.53	2.1	0.67	2.4
Real Rmse	0.32							
ADJ.SD	1.30							
Separation	4.12							
Person Reliability	.94							

Table 5: Item Separation and Reliability of assessment for learning for the faculty member's scale

	Raw Score	Count	Measure	Error	Infit		Outfit	
					MNSQ	ZSTD	MNSQ	ZSTD
Mean	95.7	30.0	.00	0.33	.99	-.3	1.14	-0.2
S.D.	5.9	0.0	0.50	0.04	0.69	2.1	1.24	2.4
Real Rmse	0.37							
ADJ.SD	0.34							
Separation	2.93							
Item Reliability	0.69							

Findings

First, to answer the first question, the instrument dimensions about the assessment for learning practices for faculty members were analyzed. The means, standard deviation, rank, and percentages of the faculty members' responses were extracted. The Likert scale has five options or values: 1,2,3,4 and 5. The means are given the following gradient based on the following equation:

$$(5-1)/5 = 0.80 \qquad (1)$$

So, the levels are shown as seen follows:

Range	Practice Degree
1.0 – 1.8	Very low
1.81 – 2.6	Low
2.61 – 3.41	Medium
3.42 – 4.22	High
4.23 – 5.0	Very high

Source: Data Adapted from the previous equation

The goal of this classification is to classify faculty members' responses. Table 6 below shows the means, standard deviation, rank, and practice degree for the faculty members on the whole scale.

Table 6: The means, standard deviation, rank, and practice degree of assessment for learning for the faculty members in the whole scale

Rank	Dimension	N	Mean	standard deviation	Practice Degree
1	AvEF	255	3.52	.61579	high
2	AvSL	255	3.42	.58853	high
3	AvEQ	255	3.41	.64957	Medium
4	AvPS	255	3.15	.59019	Medium
Overall average		255	3.38	.52795	Medium

Table 6 shows that the items scores in both dimensions effective feedback and sharing learning objectives with mean of (3.52) and (3.42) respectively, and with a standard deviation of (.61579) and (.58853) respectively. In general, the mean of both dimensions effective feedback and sharing learning objectives indicated high practice degree. While the items scores in both dimensions' effective questions and peer and self–assessment with mean of (3.41) and (3.15) respectively, and with a standard deviation of (.64957) and (.59019) respectively. In general, the mean of both dimensions' effective questions and peer and self–assessment indicated a medium practice degree. The effective feedback dimension ranked first, followed by the sharing learning objectives dimension, then the dimension of the effective question, finally the peer and self–assessment dimension.

Data analysis results also showed that the items scores in the effective feedback dimension have means ranging from (3.37) to (3.78) and standard deviation ranging from (0.71) to (2.61). Item EF4 was found to be the most practice degree and Item EF2 was found to be the lowest item in this subscale as shown in Table 7 below, while the items scores in sharing learning objectives dimension have means ranging from (3.2) to (3.6) and standard deviation ranging from (0.69) to (0.84). Item SL2 was found to be the most practice degree and Item SL7 was found to be the lowest item in this subscale as shown in Table 8 below. The items scores in the dimension of the effective question had means ranging from (3.32) to (3.54) and standard deviation ranging from (0.84) to (0.85). Item EQ2 was found to be the most practice degree and Item EQ10 was found to be the lowest item in this subscale as shown in Table 9 below. Finally, the items scores in peer and self–assessment dimension had means ranging from (2.82) to (3.46) and standard deviation ranging from (1.11) to (0.81). Item PS3 was found to be the most practice degree and Item PS4 was found to be the lowest item in this subscale as shown in Table 10 below.

Table7: The means, standard deviation, rank, and practice degree of assessment for learning for the faculty members in the second dimension (Effective Feedback)

Rank	Items	N	Mean	Std. Deviation	Practice Degree
1	EF4	255	3.7804	2.61409	High
2	EF9	255	3.5373	2.04419	High
3	EF1	255	3.5333	.75677	High
4	EF10	255	3.5020	.92589	High
5	EF5	255	3.4902	.85049	High
6	EF6	255	3.4706	.92967	High
7	EF3	255	3.4706	.93810	High
8	EF8	255	3.4667	.82162	High
9	EF7	255	3.4471	.97009	High
10	EF2	255	3.3725	.70880	Medium

Table 8: The means, standard deviation, rank, and practice degree of assessment for learning for the faculty members in the first dimension (Sharing Learning Objectives)

Rank	Items	N	Mean	Std. Deviation	Practice Degree
1	SL2	255	3.6078	.69527	High
2	SL8	255	3.4863	.68073	High
3	SL5	255	3.4863	.84579	High
4	SL3	255	3.4784	.76238	High
5	SL6	255	3.4353	.73891	High
6	SL1	255	3.4314	.98921	High
7	SL10	255	3.4196	.87867	High
8	SL4	255	3.3255	.98809	Medium
9	SL9	255	3.3059	.80876	Medium
10	SL7	255	3.1922	.83615	Medium

Table 9: The means, standard deviation, rank, and practice degree of assessment for learning for the faculty members in the third dimension (Effective Questions)

Rank	Items	N	Mean	Std. Deviation	Practice Degree
1	EQ2	255	3.5451	.84934	High
2	EQ7	255	3.4745	.87751	High
3	EQ3	255	3.4667	.81198	High
4	EQ8	255	3.4431	.79597	High
5	EQ4	255	3.4314	.91046	High
6	EQ9	255	3.3922	.81999	Medium
7	EQ6	255	3.3765	.84149	Medium
8	EQ1	255	3.3725	.99132	Medium
9	EQ5	255	3.3647	.79636	Medium
10	EQ10	255	3.3176	.84479	Medium

Table 10: The means, standard deviation, rank, and practice degree of assessment for learning for the faculty members in the fourth dimension (Peer and Self – Assessment)

Rank	Items	N	Mean	Std. Deviation	Practice Degree
1	PS3	255	3.4627	.81181	High
2	PS2	255	3.3922	.81999	Medium
3	PS6	255	3.3059	.84681	Medium
4	PS7	255	3.2745	.88466	Medium
5	PS5	255	3.1059	1.07246	Medium
6	PS9	255	3.0863	.99626	Medium
7	PS8	255	3.0510	1.01240	Medium
8	PS1	255	2.8275	1.08390	Medium
9	PS4	255	2.8235	1.11019	Medium

Answering the second question, Table 11 below showed that the means, standard deviation, and rank for practice degree of faculty members of assessment for learning. The findings were as follows: The first five items respectively EF4, SL2, EQ2, EF9, and EF1 were the most prominent the practice of assessment for learning of faculty members. While the last five items respectively ps5, ps9, ps8, ps1, and ps4 were the lowest practice of assessment for learning of faculty members.

Table 11: The means, standard deviation, rank, and practice degree of assessment for learning for the faculty members in all items of the instrument

Rank	Items	N	Mean	Std. Deviation	Practice Degree
1	EF4	255	3.7804	2.61409	High
2	SL2	255	3.6078	.69527	High
3	EQ2	255	3.5451	.84934	High
4	EF9	255	3.5373	2.04419	High
5	EF1	255	3.5333	.75677	High
6	EF10	255	3.5020	.92589	High
7	EF5	255	3.4902	.85049	High
8	SL8	255	3.4863	.68073	High
9	SL5	255	3.4863	.84579	High
10	SL3	255	3.4784	.76238	High
11	EQ7	255	3.4745	.87751	High
12	EF6	255	3.4706	.92967	High
13	EF3	255	3.4706	.93810	High
14	EF8	255	3.4667	.82162	High
15	EQ3	255	3.4667	.81198	High
16	PS3	255	3.4627	.81181	High
17	EF7	255	3.4471	.97009	High
18	EQ8	255	3.4431	.79597	High
19	SL6	255	3.4353	.73891	High
20	EQ4	255	3.4314	.91046	High
21	SL1	255	3.4314	.98921	High
22	SL10	255	3.4196	.87867	Medium
23	EQ9	255	3.3922	.81999	Medium
24	PS2	255	3.3922	.81999	Medium
25	EQ6	255	3.3765	.84149	Medium

26	EF2	255	3.3725	.70880	Medium
27	EQ1	255	3.3725	.99132	Medium
28	EQ5	255	3.3647	.79636	Medium
29	SL4	255	3.3255	.98809	Medium
30	EQ10	255	3.3176	.84479	Medium
31	SL9	255	3.3059	.80876	Medium
32	PS6	255	3.3059	.84681	Medium
33	PS7	255	3.2745	.88466	Medium
34	SL7	255	3.1922	.83615	Medium
35	PS5	255	3.1059	1.07246	Medium
36	PS9	255	3.0863	.99626	Medium
37	PS8	255	3.0510	1.01240	Medium
38	PS1	255	2.8275	1.08390	Medium
39	PS4	255	2.8235	1.11019	Medium

Answering the third question, T-Test and one-way analysis of variance was used. Table 12 below shows the results of the T-Test in the practice degree of faculty members of assessment for learning in the effective feedback, sharing learning objectives, effective questions, and peer and self–assessment dimensions due to the gender and faculty.

Table 12: The results of T-Test for differences between means according to gender and faculty

Variables and Dimensions			**No.**	**Mean**	**S. D.**	**t**	**Sig.**
Gender	AvSL	Male	171	3.4222	.58559	.207	.836
		Female	84	3.4060	.59786		
	AvEF	Male	171	3.5246	.61900	.328	.743
		Female	84	3.4976	.61252		
	AvEQ	Male	171	3.4135	.67519	.114	.909
		Female	84	3.4036	.59788		
	AvPS	Male	171	3.1076	.58732	1.512	.132
		Female	84	3.2262	.59153		
	Whole dimensions (AvTOT)	Male	171	3.3766	.53284	.231	.818
		Female	84	3.3929	.52084		
Faculty	AvSL	Humanity	177	3.4051	.58019	.481	.631
		Scientific	78	3.4436	.61000		
	AvEF	Humanity	177	3.5028	.58770	.502	.616
		Scientific	78	3.5449	.67836		
	AvEQ	Humanity	177	3.3994	.62117	.398	.691
		Scientific	78	3.4346	.71347		
	AvPS	Humanity	177	3.1254	.60545	.865	.388
		Scientific	78	3.1949	.55477		
	Whole dimensions (AvTOT)	Humanity	177	3.3672	.50593	.670	.503
		Scientific	78	3.4154	.57684		

Table 12 shows that the value of (t = 0.231) for whole dimensions indicated that there were no statistically significant differences between the means. Where the significant level is more than (0.05). In other words, there were no statistically significant differences between the responses of the sample on the practice degree of faculty members of assessment for learning based on

gender. The value of (t = 0.670) for the whole dimensions indicated that there was no statistically significant difference between the means, where the significant level is more than (0.05). In other words, there were no statistically significant differences between the responses of the sample on the practice degree of faculty members of assessment for learning according to faculty.

Table 13 below shows the results of a one-way analysis of variance in the practice degree of faculty members of assessment for learning in the effective feedback, sharing learning objectives, effective questions, and peer and self–assessment dimensions due to the academic rank and teaching experience.

Table 13: The results of analysis of variance of differences between the means of responses of sample about the practice degree of faculty members of assessment for learning

			Sum of Squares	df	Mean Square	F	Sig.
Academic Rank	AvSL	Between Groups	3.223	2	1.611	4.791	.009
		Within Groups	84.755	252	.336		
		Total	87.977	254			
	AvEF	Between Groups	2.743	2	1.371	3.693	.026
		Within Groups	93.575	252	.371		
		Total	96.317	254			
	AvEQ	Between Groups	5.623	2	2.811	6.976	.001
		Within Groups	101.551	252	.403		
		Total	107.173	254			
	AvPS	Between Groups	1.142	2	.571	1.648	.194
		Within Groups	87.332	252	.347		
		Total	88.475	254			
	AvTOT	Between Groups	2.970	2	1.485	5.518	.005
		Within Groups	67.827	252	.269		
		Total	70.797	254			
Teaching experience	AvSL	Between Groups	1.956	2	.978	2.865	.059
		Within Groups	86.022	252	.341		
		Total	87.977	254			
	AvEF	Between Groups	1.194	2	.597	1.582	.208
		Within Groups	95.123	252	.377		
		Total	96.317	254			
	AvEQ	Between Groups	1.473	2	.737	1.756	.175
		Within Groups	105.700	252	.419		
		Total	107.173	254			
	AvPS	Between Groups	1.336	2	.668	1.931	.147
		Within Groups	87.139	252	.346		
		Total	88.475	254			
	AvTOT	Between Groups	1.416	2	.708	2.571	.078
		Within Groups	69.381	252	.275		
		Total	70.797	254			

Table 13 shows that there are significant statistical differences in all dimensions of the instrument based on academic rank where the significant level is less than 0.05 meaning that there are significant statistical differences between the responses of the sample on the practice degree of faculty members of assessment for learning according to academic rank. On other hand, there were no statistically significant differences in all dimensions of an instrument based on teaching experience where the significant level was greater than (0.05). In general, there were no statistically significant differences between the responses of the sample on the practice degree of faculty members of assessment for learning according to teaching experience.

To know the direction of the differences in the academic rank, or to find out in favor of any of the four academic ranks, the Tukey test of the post-comparisons was used.

Table 14: The results of Tukey test for differences between the periods of the academic rank of faculty members of assessment for learning

Mean	(I) Rank	(J) Rank	Mean Difference (I-J)	Std. Error	Sig.
3.41	Professor	Associate Professor	-.3226	.3813	.832
		Assistant Professor	-.1694	.3317	.956
		Lecturer	.4740	.3464	.523
3.29	Associate Professor	Professor	.3226	.3813	.832
		Assistant Professor	.1532	.2746	.944
		Lecturer	.7966*	.2922	.039
3.44	Assistant Professor	Professor	.1694	.3317	.956
		Associate Professor	-.1532	.2746	.944
		Lecturer	.6434*	.2236	.026
3.62	Lecturer	Professor	-.4740	.3464	.523
		Associate Professor	-.7966*	.2922	.039
		Assistant Professor	-.6434*	.2236	.026

Table 14 shows that there were statistically significant differences between academic ranks in favor of lecturers.

Discussion of the Results

The results of the study indicated that the faculty members practice assessment for learning with a medium degree on the whole scale. These results are consistent with the study of (Alsbeeh, 2017; Al-Ahmadi, 2014; Hasan, 2012). The results are also consistent with the study of (Alsbeeh, 2017; Burns, 2010; Volante & Beckett, 2011) the effective feedback dimension was in the first rank. It is also consistent with the study of (Alsbeeh, 2017; Ryan, 2015; Almazrue, 2014; Volante & Beckett, 2011**(** that the peer and self–assessment dimension was in the last rank in the scale. However, is inconsistent with (Alshamrani, 2017, Ryan, 2015; Almazrue, 2014).

This is due to most faculty members using traditional evaluation practices, such as focusing on tests and grading, not to support student learning. As for the feedback dimension, it ranked first with high practice degree, because it is considered one of the basic skills practiced by the faculty member continuously to reach the goals. Sharing learning objectives dimension comes second with high practice degree, that is because most faculty members clarify the goals and work plan for students, by brainstorming, defining assignments, or discussing them with the required reports and information. The effective questions dimension comes third with a medium practice degree, this is because most faculty members do not listen to all students 'questions, inquiries, and their discussion with these questions. In addition, most faculties do not ask questions that stirring or motivate students' higher thinking skills or open-ended questions. This is because there is not enough time to do this. Peer and self–assessment dimensions came the last with a medium practice degree, because of the lack of efficiency and lack of educational qualification among some members of the teaching staff to use peer assessment, and their indifference to self-assessment. This type of assessment needs training to acquisition students' ability to make judgments. There is a great focus on grades and poor participation in assessment by students. In addition, the scarcity of training programs that aimed to spread assessment practices, such as tools and strategies, also focuses on the theoretical, non-applied side.

The results showed that there were no statistically significant differences between the responses of the sample on the practice degree of faculty members of assessment for learning practices according to gender. This is because the faculty members constrain the same instructions, directions, and plans. In addition, they received the same training programs. This result is consistent with the study of (Ryan, 2015; Al-Bashir and Barham, 2013; Sharah and Zaza, 2013). However, is inconsistent with the study of (Otaibi, 2018; Refaee et al., 2012; Albursan et al., 2015).

The results showed that there were no statistically significant differences between the responses of the sample on the practice degree of faculty members of assessment for learning practices according to faculty. This is because the faculty members live under one-university educational conditions. This result is consistent with the study of (Otaibi, 2018; Refaee et al., 2012). However, is inconsistent with the study of (Alshamrani, 2017; Ryan, 2015; Sharah and Zaza, 2013).

The results showed that there were statistically significant differences between the responses of the sample on the practice degree of faculty members of assessment for learning practices based on academic rank in favor of lecturers. This is because the lecturers or the assistant professors have an interest in the assessment program and process in general. This is because of their great enthusiasm, motivation, and desire for their proficiency in university teaching, which reflects positively on their satisfaction and attitudes. In addition, they were affected by the training courses that are still present in their minds.

The results also showed that there were no statistically significant differences between the responses of the sample on the practice degree of faculty members of assessment for learning practices based on teaching experience. This is because faculty members develop from themselves through courses, and cooperation between them. In addition, their possessing of assessment skills were led to reducing the variance and the differences in the number of experience years. This result is consistent with the study of (Otaibi, 2018; Alshamrani, 2017; Alsbeeh, 2017; Al-Bashir and Barham, 2013). However, is inconsistent with the study of (Ryan, 2015; Sharah and Zaza, 2013; Al-Ahmadi, 2014).

Recommendations

It is recommended to conduct training programs and workshops on strategies of assessment for learning in teaching practices in the classroom by distributing faculty members to groups, and each group contains several lecturers, assistant professors, associate professors, and professors. Also, when formulating the plan for any curriculum, it should focus on enriching the content on assessment activities, which enhance self and peer assessment skills. Furthermore, emphasizing the use of assessment for learning in the classroom because it provides some important ideas that can help both the teacher and students, and integrate them into the processes of thinking, creativity, dialogue, and decision-making, where the educational process is in the minds of both the teacher and the learner. The necessity of exchanging experiences between all faculty members, through exchanging attendance between them for lectures, or by holding weekly or monthly meetings to develop knowledge and skills about methods and strategies of assessment for learning. In addition, the relationship between students and faculty members must be good and effective, so that the teacher listens to students 'questions, discusses them, motivate them, and poses questions that stimulate students' higher skills. Training students on some assessment for learning processes is highly recommended, and conducting further studies about assessment for learning from students' viewpoint, or by using new tools such as observation card, case Study, classroom observation, personal interviews, etc. In the end, conducting further studies about the effectiveness of assessment for leaning on some cognitive and emotional variables.

Conclusion

This paper aimed to investigate and analyze the practices of assessment for learning among the faculty members at Saudi Universities. The results showed that the practice degree of assessment for learning among the faculty members was medium, where feedback dimension, ranked first with high practice degree. Sharing learning objectives dimension comes second with high practice degree, followed by the dimension of the effective question where came third with a medium practice degree. Peer and self–assessment dimensions came up the last. In addition, it also showed, there were no statistically significant differences in the practice degree of faculty members of assessment for learning according to gender, faculty, and teaching experience. While

there were statistically significant differences in all dimensions of assessment for learning based on academic rank.

References

Abdulkareem, E. & Omer, S. (2019). The Effectiveness of an Assessment for Learning Training Program in Improving the Practice of Primary School Teachers of Assessment for Learning Strategies. *Journal of Educational and Psychological Studies, 13*(2), 368-387. http://dx.doi.org/10.24200/jeps.vol13iss2pp368-387

Abdulkareem, E. & Alshaya, F. (2018). The Effect of a Proposed Model of Assessment for Learning on Sixth Grade Students' Social Skills. *Risalat Ul-Khaleej Al-Arabi, 39*(149), 57-73.

Abdulkareem, E. & Alshaya, F. (2016). A Proposed Model of Assessment for Learning Science for Elementary School Students. *International Interdisciplinary Journal of Education, 5*(12), 315-332. https://doi.org/10.12816/0035752

Akib, E. and Ghafar, M. N. A. (2019). Rasch Model Analysis of Assessment For Learning in Higher Education. *International Education Postgraduate Seminar*, Proceedings, Volume 2, 335-343.

Akib, E., & Ghafar, M. N. A. (2015). Assessment for Learning Instrumentation in Higher Education. *International Education Studies, 8*(4), 166-172. https://doi.org/10.5539/ies.v8n4p166

Al-Ahmadi, Souad. (2014). Constructive teaching practices for secondary school mathematics teachers'. *The Journal of Mathematics Education, 17*(3), 39-92.

Al-Bashir, A. & Barham, A. (2012). Using Alternative Assessment Strategies in Assessing Student's Learning in Mathematics and Arabic in Jordan. *Journal of Educational and Psychological Sciences*. *13*(1), 241-270.

Albursan, I., Alrwais, A., & Abdelfattah, F. (2015). Formative and summative evaluation practices for the intermediate mathematics teachers in the Kingdom of Saudi Arabia. *Journal of Educational & Psychological Sciences, 16*(2), 93-122.

Alkhayal, Nevine (2019). A proposed program for developing Assessment for learning (AFL) Literacy for Science Student Teachers in the Faculties of Education. *Egyptian Association for Science Education*. *22*(7), 153-212.

Almazrue, Haya. (2014). Formative assessment beliefs and practices for science teachers. *Journal of Educational Sciences, 22*(2), 279-315.

Alsbeeh, Haya. (2017). Use the assessment for learning among female mathematics teachers in secondary school in Riyadh. *Arab Foundation for Scientific Consultation and Human Resources Development. 57*(18), 1-65.

Alshamekh, Nora. (2018). *Evaluation in Education*. Alukah Network. Kingdom of Saudi Arabia.
Alshamrani, Saeed Mohammad (2017). Saudi Science Teachers' Assessment Practices in Alzelfy Educational Administration According to their Perceptions. *International Journal for Research in Education, 41*(2), 125-161.

Azizi, Issa. (2018). Effectiveness of Formative Assessment in Improving Academic Achievement among Students of Thinking and Scientific Research Skills Course, Faculty of Business Administration, Shaqra University. *Journal of College of Basic Education for educational and human sciences, 41*, 678-711.

Boone, W. J. (2016). Rasch analysis for instrument development: why, when, and how? CBE—Life Sciences Education, 15(4), rm4. https://doi.org/10.1187/cbe.16-04-0148

Burns, R. (2010). *Implementation of formative assessment strategies as perceived by high school students and teachers: Professional development implications*. Johnson & Wales University.

Chappuis, J., Stiggins, R., Chappuis, S. and Arter, J. (2011). *Classroom Assessment for Student Learning: Doing It Right, Using It Well*. Second Edition, Pearson, Assessment Training Institute, Inc.

Council of Chief State School Officers (CCSSO). (2018). Revising the Definition of Formative Assessment. Retrieved from https://ccsso.org/sites/default/files/2018-06/Revising%20the%20Definition%20of%20Formative%20Assessment.pdf

Darandari, Eqbal (2017). Learning-Oriented Assessment: Concepts, Characteristics, Strategies, Implementation, and Challenges. *Journal of Educational Sciences, 29*(1), 85-112.

Darandari, Eqbal (2014). Towards learning-oriented assessment: Best practices and Challenges. The twenty-seventh educational conference. Bahrain.

Darandari, E., & Murphy, A. (2013). Assessment of student learning in Saudi Arabia. In L. Smith & A. Abouammoh (Eds.), *Higher education in Saudi Arabia: Achievements, challenges, and opportunities.* Chapter6 (pp. 61-71). UK: Springer. https://doi.org/10.1007/978-94-007-6321-0_6

De Batisti, F.; Salini, S. & Crescentini, A. (2004): Statistical calibration of psychometric tests. Dipartimento Di Economia Politica e Aziendal. University Degli Studidi Milan.

Elrefaie, Ahmed (2011). Effect of Using Assessment for Learning (AFL) in Assessment Tasks Designing on Mathematics Achievement and Reducing Test Anxiety for First Grade High School. *Journal of Arab studies in education and Psychology*, *2*(5), 111-135.

Erwin, Akib and Mohamed Najib, Ghafar (2015). The Validity and Reliability of Assessment for Learning (AfL). *Education Journal*, *4*(2), 64-68.

Flórez, M. T., & Sammons, P. (2013). *Assessment for Learning: Effects and Impact*. CfBT Education Trust. 60 Queens Road, Reading, RG1 4BS, England.

Gardner, J. (2009). *AFL: Assessment for learning, practical guide*. Belfast; Council for the Curriculum Examination and Assessment (CCEA) Publication.

Gilles, J. L., Detroz, P., & Blais, J. G. (2011). An international online survey of the practices and perceptions of higher education professors with respect to the assessment of learning in the classroom. *Assessment & Evaluation in Higher Education*, *36*(6), 719-733.

Haroldson, R. (2012). *Student Perceptions of Formative Assessment in the Chemistry Classroom*. Dissertations. The University of Minnesota.

Hasan, I. (2012). The reality of the practice of mathematics teachers in the elementary stage of alternative assessment and relationship to their beliefs. *Arab studies in education and psychology, 29*, 171-202.

Jawawi, R; Mohammad, H; Matzin, R; Shahrill, S; Jaidin, J; Mundia, L & Shamsu, L. (2017). Engaging Students' Learning with Elements of Formative Assessment. *Turkish Online Journal of Educational Technology, 12*(1), 281-290.

Jett, P. M. (2009). Teachers valuation and implementation of formative assessment strategies in elementary science classrooms. Doctoral dissertation, University of Louisville.

Jones, N., & Saville, N. (2016). *Learning-oriented assessment: a systemic approach: 45 (studies in language testing)*. Cambridge: Cambridge University Press.

Kanjee, A & Mthembu, J. (2015). Assessment literacy of foundation phase teachers: An exploratory study. *South African Journal of Childhood Education*, *5*(1). 142- 168. https://doi.org/10.4102/sajce.v5i1.354

Keeley, P. (2015). *Science formative assessment, volume 1: 75 practical strategies for linking assessment, instruction, and learning*. Corwin Press. https://doi.org/10.2505/4/sc15_053_02_33

Mariam, N. (2016). The reality of formative evaluation among primary education teachers. Unpublished master thesis, Larbi Ben M'hidi University Oum El Bouaghi.

McCamey, R. (2014). A primer on the one-parameter Rasch model. American Journal of Economics and Business Administration, 6(4), 159. https://doi.org/10.3844/ajebasp.2014.159.163.

McDowella, L., Wakelinb, D., Montgomerya, C. & Kingb, S. (2011). Does assessment for learning make a difference? The development of an instrument to explore the student response. *Assessment & Evaluation in Higher Education, 36*(7), 749- 765.

Mofreh, S. A. M., Ghafar, M. N., Omar, A. H. H., Mosaku, M., & Ma'ruf, A. (2014). psychometric Properties on Lecturers' Beliefs on Teaching Function: Rasch Model Analysis. International Education Studies, 7(11), p47. https://doi.org/10.5539/ies.v7n11p47

Nunnally, J. C. (1994). *Psychometric Theory 3E*. Tata McGraw-Hill Education. New York.
Otaibi, A. (2018). The degree of availability of the standards of the realistic academic evaluation strategies for teaching staff viewpoints at Shaqra University. *Scientific journal for educational, psychological and social research, 37*(180), 431-467. https://doi.org/10.21608/jsrep.2018.39357

Pang, N. & Leung, Z. (2011).Teachers' competency in assessment for learning in early childhood education in Hong Kong. *Educational Research Journal, 26*(2), 199- 222.

Popham, W. & Stiggins, R. (2007). Assessing students' effect related to assessment for learning. *The Council of Chief State School Officers (CCSSO) and the FAST SCASS*, 1-8.

Refaee, A; Tawalbeh, H. and Alqaeud, I. (2012). The practicing degree of social studies teachers in Irbed Governorate to authentic assessment strategies. *Umm Al-Qura University Journal of Education and Psychology Sciences, 4*(1), 57-89.

Rizk, Fatima. (2014). The use of assessment for learning strategies in improving analytical thinking and scientific communication in science among six-grade primary school students. *Journal of Arab studies in education and Psychology. 55*(2), 141-192.

Ryan, Adel (2015). Assessment for Learning Practices (AFLP) among the Mathematics Teachers of the Basic Stage in Government Schools in Hebron From their Perspective. *The Islamic University Journal of Educational and Psychology Studies*. *23*(1), 272-300.

Sharah, N. and Zaza, D. (2013). An Investigation of the Assessment Practices of Elementary School Teachers in Jordan: Towards an Integrative and Comprehensive Model. *Journal of Educational and Psychological Sciences*, *14*(2), 73-104.

Shepard, L. A. (2000). The Role of Assessment in a Learning Culture. E*ducational Researcher*, *29*(7), 4–14. https://doi.org/10.3102/0013189X029007004

Stiggins, R. (2005). From formative assessment to assessment for learning: A path to success in standards-based schools. *Phi Delta Kappan*, *87*(4), 324-328. https://doi.org/10.1177/003172170508700414

Sumintono, B. (2018, February). Rasch Model Measurements as Tools in Assessment for Learning. In *1st International Conference on Education Innovation (ICEI 2017)*. Atlantis Press. https://doi.org/10.2991/icei-17.2018.11

Taras, M. (2010). Assessment for learning: assessing the theory and evidence. *Procedia-Social and Behavioral Sciences*, *2*(2), 3015-3022. https://doi.org/10.1016/j.sbspro.2010.03.457

Volante, L. & Beckett, D. (2011). Formative assessment and the contemporary classroom: Synergies and tensions between research and practice. *Canadian Journal of Education*, *34*(2), 239-255.

Zacharis, N. Z. (2010). Innovative assessment for learning enhancement: Issues and practices. *Contemporary Issues in Education Research (CIER)*, *3*(1), 61-70.