University Students' Perspectives Regarding Observed and Expected Learning Assessment Methods: Tool Development and National Exploration*

Majed AlKhayat **

ABSTRACT

The aims of this study were to (1) develop and validate a tool for measuring university assessment methods; (2) describe students' perspectives regarding observed and expected assessment methods (authentic, formative, diagnostic, summative); and (3) examine differences in students' perspectives regarding observed assessment methods based on their gender, academic field, and academic level. A cross-sectional descriptive survey was utilized with a sample of 1155 students. A stratified random sampling procedure was used to select a representative sample. A series of validation steps were used to examine the developed measure and an exploratory factor analysis with principal components analysis was used to examine its factor structure, which all supported the stability of the measure. Findings indicated that the most commonly used assessment method was the summative, followed by formative, diagnostic, and authentic methods, respectively. Expected assessment scores were significantly higher than observed scores. Female students had significantly higher scores on the authentic, formative, and diagnostic assessment. Students in the scientific field scored significantly higher on the authentic assessment than those in the humanities. Master students scored significantly higher than bachelors students on the authentic, formative, and diagnostic assessment. There is a need to design specialized training programs for university teachers in Jordan on each of the modern assessment methods and help them link these methods with the intended learning outcomes and effectiveness of teaching. Feedback from students should be weaved in any efforts to promote the assessment process in higher education.

Keywords: Authentic assessment; diagnostic assessment; formative assessment; summative assessment.

Introduction

Teachers, researchers and those interested in assessment acknowledge the importance of assessment in the progress of the teaching-learning process in university education (Alameri, Ismail, Akour, & Fakhouri, 2020; Karim, 2015; Tenison & Touger-Decker, 2018). Effective assessment is considered one of the pillars needed to achieve the intended learning outcomes. In fact, research has revealed that teacher's assessment practices can have powerful direct and indirect impacts on their students' performance, motivation, achievement, and preparedness to receive the educational experience in a meaningful way (Cuevas, Ntoumanis, Fernandez-Bustos, & Bartholomew, 2018). Therefore, it has been hypothesized that the quality of education can be seen through the quality of the assessment process itself (Mardapi, 2016).

Assessment was traditionally used to only confirm students' learning. Rarely was it viewed as a process of bringing out the potential that exists within students and creating an opportunity for them to demonstrate what they were able to do (Craddock & Mathias, 2019). The focus was primarily on the assessment of learning and not much on assessment for learning. University teachers currently use various assessment methods to assess students' performance and learning outcomes. Whether these methods include written or oral exams, observations, individual or group assignments, portfolios, or interviews, they should be valid, reliable, transparent, usable, comprehensive, and

* This research was funded by Deanship of Scientific Research and innovation / Al-Balqa Applied University

** Al-Balqa Applied University.

Received on 27/7/2020 and Accepted for Publication on 22/12/2020.

disciplined (Reeves, Boet, Zierler, & Kitto, 2015; Sianipar, Ansari, & Eviyanti 2018; Suharto, 2015; McCutcheon, Lohan, Traynor, & Martin, 2015).

The selection of assessment methods is often affected by multiple factors including teaching methods, educational content, educational level, and available resources (McCutcheon, Lohan, Traynor, & Martin, 2015; Regmi & Jones, 2020). Teachers are generally required to design assessment methods that correspond to such factors. The old-style assessment is based on exams given in various forms and is conducted once or several times during the school year for the purpose of obtaining information on students' achievement. Such methods are deemed ineffective as they measure simple skills and concepts that are expressed in numbers without providing valuable information about students' learning experience and without an active participation for the student in the assessment process (McCarthy, 2015). Over the last two decades, the process of assessment has been refined to become more dynamic and comprehensive, and include perspectives from students, parents, and community. Such progress required the use of various assessment strategies, models and tools for obtaining information (Cuevas, Ntoumanis, Fernandez-Bustos, & Bartholomew, 2018).

Literature on assessment has made a distinction between four assessment methods including the authentic, formative, diagnostic, and summative assessment (EPPI, 2002). *Authentic* assessment aims to assess student's ability to perform real-world tasks that demonstrate meaningful application of essential knowledge and skills. It includes tasks that are considered part of the learning process rather than just recall tests. It encourages students to use higher cognitive processes rather than reflexive thinking (Majid, 2014). Mueller (2005) defines authentic assessments as direct measures of students' acquired knowledge and skills through formal education to perform authentic tasks. The realistic contexts can make problems more engaging for students and help the teachers evaluate whether a student who can solve a problem in one context can transfer the skills to a similar setting. Pellegrino et al. (2001) assert that authentic assessments provide multiple paths to demonstration of learning in comparison to traditional assessments like answering multiple-choice questions that lack variability, owing to students' ability to demonstrate knowledge and skills they possess. Authentic tasks tend to provide more freedom to demonstrate their competencies.

Formative assessment is defined as the "process used by teachers and students during instruction that provides feedback to adjust ongoing teaching and learning to improve students' achievement of intended instructional outcomes" (Popham, 2008, p. 5). It aims to identify aspects of learning as it is developing in order to deepen and shape subsequent learning. Formative assessment provides teachers and learners with feedback in order to ensure the integrity of the learning process. If the information gathered with the assessment tools indicates unsatisfactory progress, then it is necessary to identify weaknesses and conduct remedial teaching. Satisfactory progress permits the teaching process to continue as planned. The formative assessment goes hand in hand with the teaching process; it provides both the teacher and the student with feedback; increases the class interaction process; and help teachers predict the successful performance of their students (Klinkenberg, 2017; McCarthy, 2017; Vonderwell, Liang & Alderman, 2007).

Diagnostic assessment involves the gathering and careful evaluation of detailed data to diagnose strengths and areas of need in all students in a *given* learning area. The data assist teachers to plan for appropriate pedagogy and targeted learning to more effectively scaffolding the learning needs of their students. Diagnostic assessment takes place at the beginning of a study unit in order to find a starting point, or baseline, for learning and to develop a suitable learning program. It is mainly concerned with revealing the strengths and weaknesses of the learner by using diagnostic tools for either outstanding or weak students. It also allows teachers to discover false teaching methods and try to adopt new teaching methods that are appropriate for their students (Brown, 2015).

Summative assessment involves the evaluation of participants and summarizes their progress at a specific time within a course or program through feedback and a grade (Taras, 2005). The focus of summative assessment is on the outcome of a task, such as an exam or assignment, and seeks to monitor *educational* outcomes (Shepard, 2005). Summative assessment is considered as assessment of learning rather than formative assessment which is assessment for learning. Assessment rubrics, based on a set of standards, expectations or criteria, are often used for summative assessment. These rubrics can be provided to students before they start working on the assessment task so they are

aware of the key criteria and their subsequent weighting. Summative assessments are given to students at the end of a set time period, or at the end of the semester, to assess what has been learned and how well it was learned (McCarthy, 2015). Summative assessment aims to recapitulate learning that has taken place, in order to record, mark or confirm achievements. This is the most commonly used assessment method in universities, often based on the results of the tests teachers give at the end of the month, or the middle or end of the semester to provide decisions on students' final performance (McCarthy, 2015).

Significance of the Study

Each of aforementioned assessment methods has its main objective not only to pass judgment on students' performance, but also to address students' educational problems and achieve high-level learning outcomes (Dixson & Worrell, 2016; Baleni, 2015). Unfortunately, the literature has revealed that most university teachers follow traditional methods of assessment, and a few are familiar with the authentic assessment techniques (Beziat & Coleman, 2015; Frey & Vicki, 2010; Karim, 2015). In addition, Ndalichako (2015) argued that most teachers practice with the main goal to give a degree to students after the end of the classroom learning process, and that most assessment efforts are focused on what the students learn inside the classroom only.

Effective assessment that includes aspects of authentic, formative, diagnostic, and summative assessment is needed to gain information that is reliable and consistent across classes and students in order to understand student strengths and weaknesses in relation to expected standards, and improve the overall learning outcomes (Gichuru, 2014; Mehmood et al ,. 2012).

In the Arab region, assessment methods used in higher education are based on numerical grades that get automatically and electronically converted into a symbol and a corresponding point according to a pre-set system (Alabdelwahab, 2012). Such methods rely mainly on the summative assessment, which makes the education process focused on "teaching to the test" and does not provide an accurate reflection of learning.

However, to the best of the researcher's knowledge, no studies have yet explored Arab university students' perspectives regarding the used assessment methods. In addition, there are no tools available to measure these perspectives. Given that the assessment process is considered an important indicator of the quality of education and the success of the students' learning process, can have direct impact on the students' performance and satisfaction, and can guide teachers' education process, investigating how students perceive the assessment they receive and what expectations they have is warranted.

Aims of the Study

The aims of this study were to (1) develop and validate a tool for measuring university assessment methods; (2) describe students' perspectives regarding observed and expected assessment methods (authentic, formative, diagnostic, summative); (3) examine differences between students' perspectives regarding observed and expected assessment methods; and (4) examine differences in students' perspectives regarding observed assessment methods based on their gender, academic field, and academic level. To these aims, this study targeted a national sample of Arab students from Jordan. The following section provides contextual argument relating to tertiary education in Jordan.

Brief on Higher Education Sector in Jordan

Jordan, officially the Hashemite Kingdom of Jordan is an Arab country located in Western Asia. Jordan has a population of about 10 million people. The country is disproportionately young with 35% of the population aged under 14, 61% aged 15-59, and only 4% aged 60 or above (Jordan Department of Statistics [DOS], 2019). The official language is Arabic, but English is widely spoken and is the language used in higher education institutions. UNESCO ranking for gender equality in education has put Jordan 18th worldwide (Fiske, 2018). The country strives to achieve excellence as qualified human resources who are able to compensate for the lack of natural resources. Thus, Jordanians have been oriented towards education. This has resulted in high literacy rates in the country. Illiteracy in Jordan (6.8%) is among the lowest in the region; also, 30% of men and 40% of women between 19 and 23 are attending universities.

University education is considered very important in Jordan, where people tend to perceive it as a must. This has resulted in establishing a number of public and private universities to meet the educational needs within Jordan and the region (Mahafzah, 2017).

The Jordanian higher education system follows the US model and is credit-based employing the three-cycle system: Bachelor's degrees generally require four years of study; master's degrees require one and a half to two years of study and students have the option to conclude the degree with a thesis or a comprehensive exam. Doctorate degrees require a minimum of three years of study and are only offered by a small number of selected universities. The system is regulated by two laws: the Law of Higher Education No.23 (2009) establishes the duties and powers of the Ministry of Higher Education and Scientific Research (MOHE), the Higher Education Council (HEC) and its responsibilities as well as the Scientific Research Support Fund. The Jordanian Universities Law No. 20 (2009) establishes the administrative and financial independence of public and private higher education institutions and clarifies their institutional governance (Westphal, 2019).

Higher education in Jordan plays a key role in the process of comprehensive development. Despite the limited financial and human resources, higher education lies within the priorities of the country as of the role it plays in promoting the economic, social and knowledgment level of the Jordanian citizen. During the last two decades, higher education in Jordan witnessed a significant progress in terms of the diversity of study programs ((948 for the BSc level; 476 for master's level and 105 for PhD level) and students, patterns of teaching and learning that control both the quality and quantity, and expansion of higher education institutions (Abu-El-Haija & Alkhader, 2017).

Many public and private universities were established in Jordan, this is in addition to the foreign universities operating in the country, the programs emanated from cooperation agreements between Jordan and foreign universities and the programs of the Jordanian universities in various universities of neighboring Arab countries. Currently, there are 10 public and 21 private universities in Jordan besides 51 community colleges, with around 267,000 undergraduate students, and 18,000 graduate students (MOHE, 2020). Whereas global education in the sense of independent international universities located in Jordan does not exist, global education in term of joint programs does. As a matter of fact, Jordan may be a pioneer in this field, and has responded to the increasing social demands at all levels and in all forms of higher education by the accreditation of joint program universities and new fields of study and by granting domestic universities greater autonomy. Jordan has thereby become an attractive alternative center of higher education in the region and this is clearly reflected in the number of foreign students study at the Jordanian universities which is close to 28 thousand students from around the world (Massadeh, 2020).

In spite of its successes, Jordanian higher education sector faces several challenges including the poor economy of the country, unemployment among graduates, high number of graduates and insufficient university budget, unavailable and inaccurate statistics regarding the nature of the labour market to design more efficient academic programmes, weak partnership and collaboration between private sectors and the universities, migration of well-educated individuals overseas, bureaucracy, and nonresponsive legislation (Mahafzah, 2017).

Progression of Students in Jordanian Higher Education Institutions

Departments usually develop study plans that guide students from the time they enter university until completing the requirements and obtaining a degree. A study plan guides the student for what courses to register in each semester. For obligatory courses, the courses are shown in different semesters of the study plan according to their levels. Except for the entry-level courses, each course usually has one or more prerequisites. A student may enrol in a course after successfully passing the prerequisite(s). In a couple of specialties, namely medicine and dentistry, a student must pass one year in order to continue to the next. If a student fails one or two subjects, he or she is allowed to re-examination before the beginning of the next year. Only if they pass such subjects, may they continue to the next year. Otherwise the year will need to be repeated. Besides these cases, universities do not accredit any prior experiential learning such as work, community or volunteer experience (Abu-El-Haija & Alkhader, 2017).

Student Assessment

Among the diverse range of student assessment systems and models around the world, even from high-performing education systems that range from rigorous, high-stakes and frequent testing systems such as those in China, Korea, and Vietnam in Asia, to more flexible and low-stakes and infrequent testing systems such as those in Finland, the Netherlands, and Slovenia in Europe, Jordan, stands in the "middle" in terms of testing frequency, types of student assessments, grade levels, subject domains, as well as testing instruments, rubrics development, and administration (Ababneh et al., 2014).

However, most of applied assessment methods depends on exams. Usually each course requires students to sit for at least two semester exams and one final exam. Laboratory courses may require students to sit for one written exam and take one practical (oral) exam. Semester work (exams plus any other assignments, projects, etc.) makes 60% and the final exam makes up 40% of the overall grade. In order to pass a course, a student must obtain 50% or more. The grade point average for students to remain in good standing (i.e. not to be put under probation) and to graduate upon completion of all requirements is 60%. If the cumulative grade of a student falls below these thresholds, they receive a written warning asking him/her to raise a grade within a given period. If the student fails to do that, then they are given a final warning and if they still fail to raise the grade, the student will be dismissed from the university (department or faculty) or the college (Abu-El-Haija & Alkhader, 2017).

E-Learning in Higher Education

The Jordanian higher education system has proven to be among the fastest-growing educational systems in the Arab district system (Alkhawaja & Abd Halim, 2019). E-learning was first introduced in 2002, opening the gates to technological usage for education in students and faculty members of various Jordan universities that require pedagogical and instructional support as well as improvements (Algahtani, 2016). The Ministry of Education of Jordan and the ministries of planning and information technology and telecommunications coordinated towards fulfilling national e-learning through the creation of national knowledge networks. This e-learning system relies on the development of self-learning and critical thinking instead of the traditional educational methods, which heavily depend on indoctrination by the teacher or lecturer (Dirani & Yoon, 2009).

E-learning offers alternative approaches to Jordanian traditional higher education institutions, encouraging them to re-evaluate the way they operate. In doing so, it provides potential to accommodate new information and communication technologies to enhance the student learning experience (Al-Adwan & Smedley, 2012). However, Al-shammary and Ali (2017) emphasized that even though the communication technology is available and well established in the region in general and Jordan in particular, its use in higher education is still relatively limited. This limitation has been inferred to cost of development and maintenance (Bringula & Basa, 2011), faculty and students attitudes towards elearning as many possess strong allegiances towards the traditional teaching model (Al-azawei, Parslow, & Lundqvist, 2017), lack of incentives and rewards systems in encouraging faculty participation, lack of institutional staff training (Al-Shboul, 2013), lack of legal frameworks and censorship systems, and weak infrastructure in the adoption and quality improvement of e-learning (Osman, 2018), students' lack of necessary hardware and skills (Al-Adwan & Smedley, 2012), and changes in the cultural expectations (Al-Adwan & Smedley, 2012). Moreover, educational systems in Jordan are overstressed by the flow of Arab refugees, creating challenges for the adoption of new techniques and strategies to provide quality education for all (Jalbout & Farah, 2016).

Another important issue to highlight is that socio-cultural factors may pose several barriers during the implementation of virtual learning. For instance, eye-to-eye contact, especially between males and females in many different Arab countries, is deemed contrary to family traditions, which encourages humility (Rhema & Mlliszewska, 2010). A study on factors that affect implementing e-learning in the Jordanian higher education system (Al-Adwan & Smedley, 2012) revealed that working independently was unpopular among university students, with the majority indicating that face-to-face contact with tutors was a vital part of their learning. They strongly believed traditional classes helped them to understand the content of lectures. This phenomenon is considered an essential component of

the Jordanian education system where teachers are the center of the educational process. Students in Jordan are used to strictly following tutors' directions with teachers seen as the main source of motivation and information for students. Additionally, many students in Jordan consider the internet as a device for passing time and communicating with others. Within such an educational culture, students' attitude is negatively positioned towards e-learning as they defer to their teachers' direct instructions rather than following independent thinking (Al-Adwan & Smedley, 2012).

Methods

Design

A cross-sectional descriptive survey was utilized. Data were collected using a self-administered questionnaire that included measures on assessment methods and students' characteristics. A stratified random sampling procedure was used to select a representative sample from all schools at Al-Balqa Applied University in accordance to the formula nh = (Nh / N) * n. Al-Balqa University was selected for this study because it is one of the largest universities in Jordan and has campuses distributed in all regions of the country. The survey was designed based on lessons learned from an earlier pilot study. The pilot study was conducted to guide the planning of the large-scale investigation with aims to assess recruitment potential; assess the feasibility of collaboration and coordination with the proposed recruitment settings; increase clinical experience with the study protocol; and identify any potential obstacles and challenges in a reasonable period of time. A total of 30 participants were included in the pilot study. Overall evidence of feasibility of the study was identified and the goals of piloting were successfully achieved. The study obtained Institutional Review Board approvals from Al-Balqa Applied University.

Sample Size and Power Estimates

The minimum sample size needed for this study was calculated using the following parameters: (a) a precision rate of 3%; (b) a population size of 33,456; and (c) a 95% confidence interval (CI). The calculated minimal sample size needed for the study was 1,035.

Measure

Tool Development and Validation

No tools are currently available to measure students' perspectives regarding assessment methods used by their faculty. Therefore, a tool was developed for the purposes of this study. Several steps were taken to develop this tool, named The Assessment Methods Questionnaire.

First: A systematic review of the literature was conducted to locate studies in the psychological, cognitive and educational fields that include relevant data on assessment methods and techniques (e.g., Frey & Vicki, 2010; Thorndike, 2005; Zhang & Stock, 2003). The investigator (AA) and two research assistants conducted the systematic review of several database using search algorithms to identify relevant studies. The two research assistants were qualified educators who hold master's degrees in psychology and have experience in research and literature review. An academic health center reference librarian helped building a combination of index and Medical Subject Headings (MeSH) terms, which was used according to the requirements of each database. The researcher, who is a native Arabic speaker, conducted another search using related Arabic terms using different Arabic databases. No restrictions were applied on publication status and date. Articles were required to feature keywords in the title and/or abstract of full-length publications, be peer-reviewed, and be written in English or Arabic. To ensure a quality search, the investigators adhered to the Preferred Reporting Items for Systematic Reviews statement (Moher, Liberati, Tetzlaff, & Altman, 2009) to guide the process of identification, selection, and appraisal of the reviewed articles. Each potential article was evaluated based on quality criteria extracted from the WHO Reproductive Health Library guidelines for critical appraisal of systematic reviews (Abalos, Carroli, Mackey, Bergel, & de Estudios Perinatales, 2001) and Meline's (2006)inclusion and exclusion criteria for selecting studies for systematic review. A structured study evaluation form was developed based on the previous criteria and was independently used by the three reviewers to evaluate the elicited studies. The goal of the form was to minimize bias in locating, selecting, coding, and aggregating individual studies.

Second: Data retrieved from step one was used to build individual items that measure four assessment methods

(authentic, formative, diagnostic, and summative). Students' perspectives on each item was measured as an observed and expected method. A panel of three experts holding PhD in education and have extensive experience in tools development and validation was responsible for creating the individual items. All members were familiar with the terminology of the area covered by the measure. Together, the panel reviewed the items to identify and resolve any inadequate expressions/concepts as well as any discrepancies between the developed items. The panel also suggested alternative words or expressions as needed. Items were carefully written to meet the following criteria: Be simple and direct avoiding jargons and unnecessary wording; be written in a positive format; starts with a verb that reflects the action being evaluated; and be free of any type of biases. After reviewing the final edits, the panel agreed on a final version of the measure.

Third: Face validity was conducted and concluded that the tool measures the characteristic or trait of interest. The tool was sent to experts in the field to look into the items in the questionnaire and provide feedback on whether the tool was a valid measure of the concept which is being measured. In particular, those experts were required to evaluate whether each of the measuring items matches the given conceptual domain of the concept.

Fourth: A Principal component analysis (PCA) with varimax rotation was conducted with 57 items of the assessment methods questionnaire. Based on the following findings, data deemed to be appropriate for conducting the PCA included: (a) the correlation matrix determinant = .194; (b) significant Bartlett's test of sphericity ($p < .0001$); and (c) KMO value = .76. According to Nunnally and Bernstein (1994), determining the number of factors to be extracted depends on how strongly and cleanly the variables load on the factors. The variable loads strongly on a particular factor if loading ≥ .40 and is considered clean if the absolute difference between the loadings is more than .20. Accordingly, 4 items were deleted leaving a final total of 53 items. The PCA yielded four distinct factors that explained almost

54% of the total variance. These factors were described as the following: authentic assessment; formative assessment; diagnostic assessment; and summative assessment. Correlation coefficients for the four factors were as the following: 1*2=0.342; 1*3=0.446; 2*3=0.266; 1*4=0.252; 2*4=0.411; 3*4=0.311.

Fifth: Pilot testing to refine or modify the measure as needed was conducted with a sample of 30 students. The questionnaire was distributed for the pilot sample to fill then a focus group discussion was utilized to gain feedback from students regarding the clarity of the concepts. Several linguistic modifications were conducted based on the pilot results.

Tool Description

The Assessment Methods Questionnaire that was used in this study to measure university students' perspectives regarding observed and expected assessment methods is a self-reported questionnaire of 53 items, each of which consists of four self-evaluative statements scored 1 (unused) to 5 (much used). Higher total scores indicate greater use of the corresponding assessment method. The measure includes four subscales representing four assessment methods: authentic (17 items, score range: 17 – 85); formative (14 items, score range: 14-70); diagnostic (12 items, score range: 12-60); and summative (10 items, score range: 10-50). Using the sample of 1155, the scale and subscales showed adequate reliability with the following Cronbach's alpha scores: Total scale= .85; authentic= .73; formative= .70; diagnostic= .75; and summative= .55).

Data Analysis

The collected data were entered and analyzed using the Statistical Package for Social Sciences (IBM, SPSS, 21). Many steps were conducted to secure reliable and valid statistical results. Firstly, the type of measurement level for each variable was defined to help deciding how to organize and display data. A thorough univariate analysis was conducted. Data were inspected for outliers, defined as values greater than 1.5 standard deviation units from the sample mean for a given variable (Polit, 2009). Data were also checked for possible missing entries and those were handled using the method described by Munro (2005). First, the type of measurement scale for the variable was defined to help deciding how to replace its missing data. Second, the 'amount' of missing was checked. Any questionnaire with more than 50% missing data was discarded. In addition, the 'pattern' of the missing; whether it has a systematic or random

fashion throughout the data set, was determined. Third, the underlying reason that would contribute to missing was decided or estimated. Fourth, a method of 'estimating missing data though imputation' was used. Particularly, through replacing missing data by the mean of the scores (or the median if the data are categorical, or continuous with a skewed distribution). Fifth, any proposed inferential statistical tests were run twice, first before replacement of the missing data and then after the replacement to make sure that the method of imputation is appropriate (Musil et al., 2002).

Finally, non-directional statistical tests were conducted with the level of significance set at 0.05 for each test. Due to the exploratory nature of the study, the level of significance was not adjusted for multiple tests. To answer the research questions, the following statistical tests were run: (1) descriptive statistics to summarize sample characteristics as well as the study scale; (2) paired t-test to examine differences between students' perspectives regarding observed and expected assessment methods; and (3) independent sample t-test to test differences in students' perspectives regarding observed assessment methods (authentic, formative, diagnostic, summative) based on their gender, academic field, and academic level.

Results

Participants

A total number 983 undergraduate students and 349 graduate students returned the study survey. However, almost 6% of the questionnaires collected from each campus were either empty or had more than 50% of the questionnaire data missing, and therefore were excluded. The final sample included in the analysis was 1155 students (55% males; 75% undergraduates). Table 1 shows the recruited sample distribution.

Table 1. ***Study sample distribution according to academic field and level (including excluded questionnaires)***

Faculty (Bachelor's)	Bachelors			Total
	Regular	Parallel	International	
Main Campus	335	53	1	388
Faculty of Engineering Technology	114	38	5	157
Princess Alia University College	68	7	1	75
Amman College for Financial Science	56	5	0	62
Zarqa University College	34	3	0	38
Princess Rahma University College	20	0	0	21
Husson University College	72	6	1	79
Ajloun University College	72	1	0	74
Irbid University College	71	3	0	73
Karak University College	1	0	0	2
Shoubk University College	5	1	0	5
Maan College	0	0	0	0
Aqaba University College	10	1	0	11
Total	859	117	7	983
Master' Students				394

Aim 1. ***Describe students' perspectives regarding observed and expected assessment methods (authentic, formative, diagnostic, summative).***

The mean response scores of students' perceptions regarding *observed* assessment methods were 2.9 (±.4) on the authentic assessment scale; 3.4 (±.6) on the formative assessment scale; 3.2 (±.5) on the diagnostic assessment scale; and 3.5 (±.6) on the summative assessment scale. The mean response scores of students' perceptions regarding

expected assessment methods were 1.7 (±.4) on the authentic assessment scale; 4.3 (±.3) on the formative assessment scale; 4.4 (±.3) on the diagnostic assessment scale; and 4.8 (±.4) on the summative assessment scale. Figure (1) displays the responses means for observed and expected assessment methods.

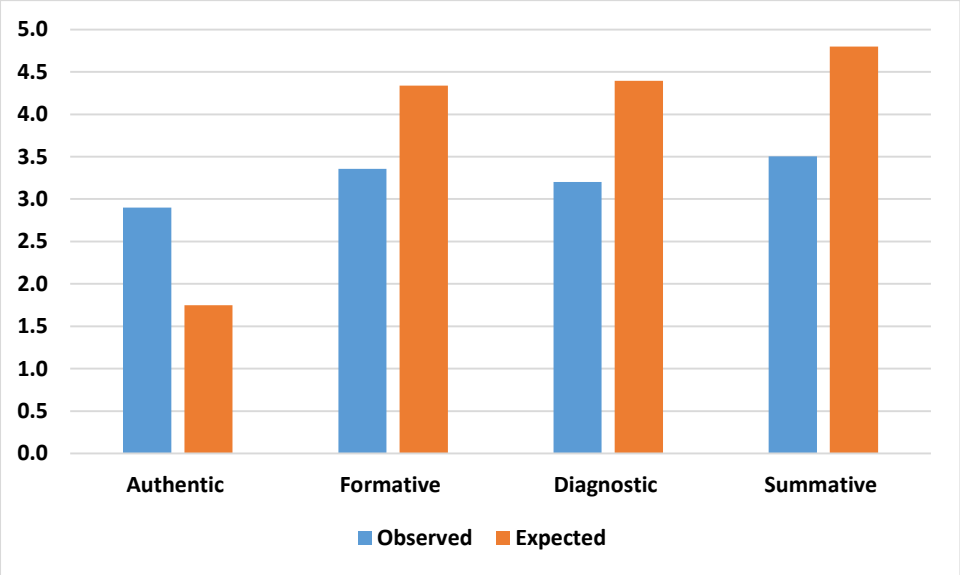

Figure 1. Mean scores for the four assessment strategies

Aim 2. ***Examine differences between students' perspectives regarding observed and expected assessment methods.***

Paired t-test was used to examine differences between students' perspectives regarding observed and expected assessment methods (authentic, formative, diagnostic, summative). The results revealed that expected formative, diagnostic, and summative assessment scores were significantly higher than observed scores. On the other hand, expected authentic scores were significantly lower than the observed (Table 2).

Table 2: D*ifferences between students' perspectives regarding observed and expected assessment strategies (N=1155)*

Assessment Method		**t**	***df***	***p***
Authentic	Observed	88.870	1154	<.001*
	Expected			
Formative	Observed	-52.941	1154	<.001*
	Expected			
Diagnostic	Observed	-58.985	1154	<.001*
	Expected			
Summative	Observed	-52.059	1154	.001*
	Expected			

* *p* less than .05, statistically significant

Aim 3: ***Examine differences in students' perspectives regarding observed assessment methods (authentic, formative, diagnostic, summative) based on their gender, academic field, and academic level.***

Three independent sample t-test were run to examine differences in students' perspectives regarding observed assessment methods based on their gender, academic field, and academic level. The analysis showed significant gender differences in the authentic (t (1154) = 4.9, p< .001), formative (t (1154) = 5.3, p< .001), and diagnostic assessment (t (1154) = 6.8, p< .001) with females scoring higher than their male counterparts (Table 3). No significant gender difference was found for the summative assessment (p>.05).

Table 3. ***Differences in students' perspectives regarding observed assessment strategies based on their gender (Males=633; Females= 522)***

Assessment Strategy		Mean	SD	t df	*p*
Authentic	Female	3.3	.4	4.9 1153	.005*
	Male	3.1	.4		
Formative	Female	3.4	.6	5.3 1153	.002*
	Male	3.2	.6		
Diagnostic	Female	3.5	.4	6.8 1153	.001*
	Male	3.3	.5		
Summative	Female	3.5	.5	2.1 1153	.06
	Male	3.4	.6		

* p less than .05, statistically significant

With regard to students' academic field (scientific, humanities), there were significant differences only in the authentic assessment (t (1153) = -14.37, p=.001) with students in the scientific field scoring higher, and in the formative assessment (t (1153) = -5.77, p< .001), with students in the humanities field scoring higher (Table 4).

Table 4. ***Differences in students' perspectives regarding applied assessment strategies based on their academic field (Scientific =699; Humanities = 456)***

Assessment Strategy		Mean	SD	t df	*p*
Authentic	Scientific	3.4	.5	14.3 1153	.001*
	Humanities	3.1	.1		
Formative	Scientific	3.2	.6	-5.7 1153	.001*
	Humanities	3.4	.2		
Diagnostic	Scientific	3.4	.6	-2.1 1153	.06
	Humanities	3.5	.2		
Summative	Scientific	3.4	.7	-1.4 1153	.14
	Humanities	3.5	.3		

* p less than .05, statistically significant

With regard to students' academic level (bachelor, master), the analysis showed significant differences in the authentic (t (1153) = -3.3, p=.001), formative (t (1153) = -4.2, p< .001), and diagnostic assessment (t (1153) = -2.4, p< .02) with the master's students scoring higher than bachelor's students (Table 5). No significant difference was found regarding the summative assessment (p>.05).

Table 5. ***Differences in students' perspectives regarding applied assessment strategies based on their academic level (bachelor =864; master = 291)***

Assessment Strategy		Mean	SD	t_{df}	*p*
Authentic	Bachelor	3.1	.4	-3.3_{1153}	.001*
	Master	3.4	.3		
Formative	Bachelor	3.2	.5	-4.2_{1153}	<.001*
	Master	3.4	.6		
Diagnostic	Bachelor	3.4	.5	-2.3_{1153}	.01
	Master	3.5	.4		
Summative	Bachelor	3.4	.6	$.3_{1153}$	.76
	Master	3.5	.5		

* *p* less than .05, statistically significant

Discussion

This study aimed to develop a tool for measuring assessment methods, describe students' perspectives regarding observed and expected assessment methods (authentic, formative, diagnostic, summative), examine differences between students' perspectives regarding observed and expected assessment methods, and examine differences in students' perspectives regarding observed assessment methods based on their gender, academic field, and academic level. The study introduced the first tool to measure students' perspectives regarding assessment methods. Findings proved the measure as reliable and valid. Yet, further studies are needed to replicate the measure's psychometrics across different samples.

According to the study participants, the most commonly used assessment method was the summative, followed by formative, diagnostic, and authentic methods, respectively. These findings came in line with previous studies that showed more reliance on summative assessment methods (Beziat & Coleman, 2015; Frey & Vicki, 2010; Karim, 2015). Authentic assessment method was the least reported, meaning that university faculty do not often provide students with real-world tasks that demonstrate meaningful application of essential knowledge and skills. This finding warrants careful attention from higher education policy makers and educators. In order to better prepare students to be productive citizens, they must be capable of performing meaningful tasks that replicate real world challenges. Therefore, schools must help students become proficient at performing the tasks they will encounter when they graduate. When doing so, it is important to realize that authentic assessment needs to drive the curriculum. That is, teachers first determine the tasks that students will perform to demonstrate their mastery, and then a curriculum is developed that will enable students to perform those tasks well, which would include the acquisition of essential knowledge and skills (Botelho, Agrawal, & Bornstein, 2019; McDonald, 2002).

Expected formative, diagnostic, and summative assessment scores were significantly higher than observed scores. These results indicate that students have a clear interest in these methods, and they aspire to have them used by their teachers. On the other hand, expected authentic scores were significantly lower than the observed. The latter result came against the expectation. It is possible that students were not familiar with authentic assessment and are not accustomed to application. Therefore, they did not expect using such methods to assess their performance.

Female students had significantly higher scores on the authentic, formative, and diagnostic assessment. This result indicates that female students' perception of the classroom assessment process differs from male students, and it comes in line with Ndalichako's (2015) study which found that female students are more often concerned with the assessment methods than their male counterparts. The summative traditional methods though seemed to have same perceptions by both genders.

The study also revealed that students in the scientific field scored significantly higher on the authentic assessment than those in the humanities, while the opposite revealed for the formative assessment. In addition, master students

scored significantly higher than bachelors students on the authentic, formative, and diagnostic assessment. These results might be related to the nature of the educational content, which tends to be relatively more pragmatic in the scientific schools than in the humanities. Master students also often receive variety of assessment methods compared to bachelor students and their expectation regarding the use of authentic methods are often higher.

Study Limitations

This study shed light on important areas that need strengthening. However, before presenting the potential implications and recommendations, findings need to be considered within the context of their limitations. The cross-sectional descriptive design of this study may limit the explanation of the causal relationship between the variables. Also, the study utilized a self-reported questionnaire. Reliability of the opinion expressed by the participant is an issue as there is no guarantee of accuracy. Self-selection bias is another limitation of survey research. Finally, there is still a need to consider an array of variables related to student perception regarding assessment methods to provide a clearer vision on this topic.

Implications

Assessment is an essential part of the teaching-learning process that requires careful consideration by educators, policy makers, and students themselves. Assessment should be viewed as a dynamic process rather than a judgment on students' performance. Assessment affects decisions about grades, placement, advancement, instructional needs, curriculum, and, in some cases, funding. Whether teaching at the undergraduate or graduate level, it is important for instructors to strategically assess the effectiveness of their teaching by measuring the extent to which students in the classroom are learning the course material. Designing assessment methods should start before teaching takes place and continues during and after teaching ends.

The study highlight the need to design specialized training programs for university teachers in Jordan on each of the modern assessment methods and help them link these methods with the intended learning outcomes and effectiveness of teaching. Teachers in return need to familiarize students with the used assessment method and help them appreciate the assessment process as part of their learning and success process rather than a teacher judgment on their performance.

There is a need to consolidate the current loosely-coupled student assessment systems into a well-integrated holistic student assessment system that includes various forms and designs of assessment tools for the purpose of improving equitable learning and achievement. Jordan should capitalize on the work already done in the area of student assessments so far and utilizes the existing local capacity (Ababneh et al., 2014).

A standardized testing system, particularly multiple-choice testing, may not align well with the goal of developing a set of new competencies to succeed in the global economy and society (Harvard Education Letter 2013). A well-integrated student assessment system that includes various forms and designs of assessment tools for the purpose of improving equitable learning and achievement is imperative in Jordan.

Jordan is facing more than ever the growing challenges to enhance the quality, relevancy, analytics, usefulness, and trustworthy of the existing assessment systems. Some relevant questions have already been surfaced such as: 1) How to integrate various assessment tools into strategically organized and effective educational assessment system? 2) How to design and upgrade an assessment tool within the educational assessment system that is locally curriculum relevant? 3) How to provide diverse types of succinct and useful assessment information or results? 4) How to gain the public trust when the student assessment data is released? Answers to these questions require a strategic vision and thinking for a more integrated system of student assessments in Jordan.

Jordan's education is facing a critical and historical moment of globalization, technology advance and geo-political changes. Jordan must have strong and sustainable technical forces and institutions to continuously develop and upgrade the assessment systems. It is educators' responsibility to meet the challenges of the dynamic education world (Ababneh et al., 2014).

The potential *cultural influences* on the assessment system needs to be carefully considered. According to Vrazalic

et al. (2009), culture and traditions are strongly linked to acceptable learning practices. Accordingly, specific styles of e-learning could be very successful in certain cultures but totally rejected in others. As mentioned in the introduction section, assessment modalities that require eye-to-eye contact via online outlets might not be acceptable (Rhema & Mlliszewska, 2010). Further, assessment strategies that require independent learning is not popular among students in Jordan where teachers are the center of the educational process (Al-Adwan & Smedley, 2012). These factors need to be considered when designing a successful assessment strategies.

Authenticity has been identified as a key characteristic of assessment design, which promotes learning. It encourages the learner to integrate knowledge and skills, and act on knowledge. Research has identified several positive impacts of authenticity in learning on student learning such as improving problems solving skills, autonomy, motivation, self-regulation and metacognition (Majid, 2014; Mueller, 2005; Pellegrino et al., 2001). Authentic assessment encompasses the process as well as the end-product, giving students experience of working in teams and in ways they are likely to meet in real-world tasks. While the introduction of authentic assessment is warranted, it should be noted that it has significant barriers. Of those is the lack of conceptualization of the term authentic assessment sufficient to inform assessment design at the individual course (Villarroel et al., 2018). In addition, it has been argued that teachers are often reluctant to change formal assessments, such as exams, as change require added demands on time, energy, and intellectual resources (Deeley & Bovil, 2017). There is also the challenge that such change requires teachers to have deep disciplinary knowledge and sufficient cognitive flexibility to monitor, challenge, and guide learners toward problem solutions (Saye, 2013). Policy makers and educators of need to consider these factors prior to implementing major changes in the assessment methods.

Higher Education Assessment in the Era of COVID-19

The global disruption caused by the COVID-19 pandemic has affected day to day life and unsettled many services. In higher education, the impact was swift and profound, as educational institutions had to cope with unprecedented closures, distance learning and rising issues related to academic integrity (Quacquarelli Symonds, 2020). Studies conducted during the COVID-19 period found that students worldwide have been dramatically affected by the spread of COVID-19 due to the impact that closure of the educational institutions had on their lifestyles, and reported lack of motivation and negative attitudes towards learning online (Killan, 2020; Quacquarelli Symonds, 2020).

The assessment process is considered a key challenge during these times, as most of the traditional ways of testing became no longer feasible or appropriate for large cohorts. There is obviously a need for well-designed digital assessment systems that can help achieving the intended learning outcomes under such global crises. In their report on the 'future of assessment', Pauli and Ferrell (2020) demonstrated that effective digital assessment needs to be (a) relevant by enabling universities to go beyond traditional forms of assessment, (b) adaptable by addressing the needs of a growing and diverse student population, (c) trustworthy by having solid foundations of academic integrity, security, privacy and fairness, (d) continuous by reflecting the fact that students today need to adapt to changes in the world of work rather than aiming to succeed at certain high-stress exams, and (e) secure by ensuring that the right student is taking the right assessment and that the work they are submitting is their own and abides by the rules. Taken together, these principles offer a holistic approach to more effective assessment.

This study recommended more integration of authentic assessment approaches in higher education. However, to apply this in the digital age, there is a need to understand 'who' we are assessing and their stage of study (bachelors, graduates, scientific/humanities/health fields students). This enables a better understanding of the short- and long-term consequences of decisions taken around assessment (Norcini et al. 2018). There is also a need to understand 'what' we are assessing. Assessment of clinical, communication, and practical skills cannot be replaced by formats that do not require an examiner to observe the student's performance (Boursicot et al. 2020). There is evidence that the use of some advanced technologies, simulations and videos can meet such goal (Yeates et al., 2019), both at undergraduate (Waseh & Dicker, 2019) and postgraduate levels (Lee & Nambudiri, 2019). The need to improve currently available tech and inspire a faster transformation has become more obvious if to create long-term changes to the assessment

process. By doing so, higher education institutions can prepare students for what they are going to do next, and test their knowledge and skills in a more realistic, contextualized and motivating way. In fact, technology can offer assessment opportunities to test knowledge and skills in a more realistic and motivating way than pen and paper tests, which can appear irrelevant outside the academic world.

The major lesson learned by higher educational institutions affected by COVID-19 might be the need to go beyond setting up temporary intervention alternatives in crisis to setting long-term strategies that are adaptable to changes as they come (Duraku & Hoxha, 2020). Educational institutions need to focus on improving online learning, integrating technology into teaching and assessment processes, and ensuring equal opportunities for all students to receive affordable and sustainable educational solutions (IESALC, 2020; Mintz, 2020).

Conclusions

This study has introduced a new tool for researchers and educators to measure students' perspectives regarding observed and expected assessment methods. This tool will aid exploring a major part of the learning process, assessment, from a students' perspective, which may have direct impact on the students' performance and satisfaction. While this study provided insights on students' perspectives regarding observed and expected assessment methods, there is still a need for qualitative studies to capture the more nuanced and subjective aspects of students' knowledge and attitudes towards different assessment methods. Feedback from students should be weaved in any efforts to promote the assessment process in higher education.

In the post-COVID-19 period, providing effective balance between varieties of assessment methods is crucial. Done improperly, it can be a source of dissatisfaction, frustration and anxiety. Is it valid? Is it reliable? Is it susceptible to cheating? Does it involve a sustainable workload? What logistics it needs? The COVID-19 crisis has shed light on these questions and others. Organized efforts are needed to make assessment smarter, faster, fairer and more effective.

REFERENCES

Ababneh, A., Ababneh, I., Lebdih, K. A., & Tweissi, A. (2014). Mapping of Student Assessments in Jordan. Retrieved from http://www.nchrd.gov.jo/assets/PDF/Studies/En/Mapping%20of%20Student%20Assessments%20in%20Jordan%20Sept%202014%20(2).pdf

Abu-El-Haija, A. & Alkhader, R. (2017). Overview of the Higher Education System in Jordan. European Union. Retrieved from https://eacea.ec.europa.eu/sites/eacea-site/files/countryfiches_jordan_2017.pdf

Alabdelwahab, S. Q. (2012). Portfolio assessment: A qualitative investigation of portfolio self-assessment practices in an intermediate EFL classroom, Saudi Arabia (Doctoral dissertation, The Ohio State University).

Al-Adwan, A., & Smedley, J. (2012). Implementing e-learning in the Jordanian Higher Education System: Factors affecting impact. *International Journal of Education and Development using ICT*, *8*(1).

Alameri, J., Ismail, H. B., Akour, A., & Fakhouri, H. N. (2020). Blended Learning and the Use of ICT Technology Perceptions Among University of Jordan Students. *Information, Communication and Computing Technology* (pp. 271-280).

Al-Azawei, A., Parslow, P., & Lundqvist, K. (2017). Investigating the effect of learning styles in a blended e-learning system : An extension of the technology acceptance model (TAM), *33*(November), 1–23. https://doi.org/10.14742/ajet.2758

Al-Gahtani, S. S. (2016). Empirical investigation of e-learning acceptance and assimilation : A structural equation model. *Applied Computing And Informatics*. https://doi.org/10.1016/j.aci.2014.09.001

Alkhawaja, M. I., & Abd Halim, M. S. B. (2019). Challenges of E-Learning System Adoption in Jordan Higher Education. *International Journal Of Academic Research In Business And Social Sciences*, *9*(9).

Al-Shammary, S., & Ali, A. A. (2017). ICT Hindering Factors Applied in Jordan Construction Projects. *Civil Engineering and Architecture 5(3):*, *5*(3), 83–88. https://doi.org/10.13189/cea.2017.050302

Al-Shboul, M. (2013). The level of E-learning integration at the University of Jordan: Challenges and opportunities.

International Education Studies, *6*(4), 93–113. https://doi.org/10.5539/ies.v6n4p93

Baleni, Z. G. (2015). Online formative assessment in higher education: Its pros and cons. *Electronic Journal of e-Learning*, *13*(4), 228-236.

Beziat, T. L. R., & Coleman, B. K. (2015). Classroom assessment literacy: Evaluating pre-service teachers. The Researcher, 27(1), 25-30.

Botelho, M. G., Agrawal, K. R., & Bornstein, M. M. (2019). An systematic review of e-learning outcomes in undergraduate dental radiology curricula—levels of learning and implications for researchers and curriculum planners. *Dentomaxillofacial Radiology*, *48*(1), 20180027.

Boursicot, K., Kemp, S., Ong, T. H., Wijaya, L., Goh, S. H., Freeman, K., & Curran, I. (2020). Conducting a high-stakes OSCE in a COVID-19 environment. *MedEdPublish*, *9*.

Bringula, R. P., & Basa, R. S. (2011). Factors affecting faculty web portal usability. *Educational Technology and Society*, *14*(4), 253–265.

Brown, J. D. (2015). *Testing in Language Programs*. New York: McGraw-Hill.

Craddock, D., & Mathias, H. (2009). Assessment options in higher education. Assessment & Evaluation in Higher Education, 34(2), 127-140.

Cuevas, R., Ntoumanis, N., Fernandez-Bustos, J. G., & Bartholomew, K. (2018). Does teacher assessment based on student performance predict motivation, well-being, and ill-being?. Journal of school psychology, *68*, 154-162.

Deeley, S.J. & Bovill, C. 2017. "Staff student partnership in assessment: enhancing assessment literacy through democratic processes". A*ssessment & evaluation in Higher Education* 42 (3): 463-477.

Dirani, K. and Yoon, S. (2009). *Exploring Open Distance Learning at a Jordanian University: A Case Study*. The international review of research in open and distance learning, 10 (2), 1-6.

Dixson, D. D., & Worrell, F. C. (2016). Formative and summative assessment in the classroom. Theory into practice, *55*(2), 153-159.

Duraku, Z., & Hoxha, L. (2020). The impact of COVID-19 on higher education: A study of interaction among students' mental health, attitudes toward online learning, study skills, and changes in students' life.

Fiske, E. B. (2018). *World atlas of gender equality in education*. Unesco.

Frey, Bruce B. & Schmitt, Vicki L. (2010). Teachers' Classroom Assessment Practices. Middle Grades Research Journal, 5(3), 107-117.

Gichuru, Francis Maina. (2014). Classroom assessment practices in Kenyan secondary schools: teacher perspective, Published on The Department of Psychology (http://psychology.uonbi.ac.ke).

IESALC (2020). COVID-19 and higher education: Today and tomorrow. Impact analysis, policy responses and recommendations. Retrieved from http://www.iesalc.unesco.org/en/wp-content/uploads/2020/04/COVID-19-EN-090420-2.pdf Accessed 9 September 2020

Jalbout, M., & Farah, S. (2016). *Will the Technology Disruption Widen or Close the Skills Gap in the Middle East and North Africa*. March.

Jordan Department of Statistics [DOS] (2019). Estimated population of Jordan. Retrieved from https://knoema.com/atlas/sources/DOS-JO?topic=Demographics

Killian, J. (2020). College students, professors, adjust to COVID-19 life. NC Policy Watch, 1.

Lee, M. S., & Nambudiri, V. (2019). Integrating telemedicine into training: adding value to graduate medical education through electronic consultations. *Journal of graduate medical education*, *11*(3), 251-254.

Mahafzah, (2017). Higher education in Jordan: history, present status and future. QS Global Education News. Retrieved from https://qswownews.com/higher-education-in-jordan/

Majid, A 2014, Implementation of Curriculum 2013 Theoretical Study and Practical, Interest Media, Bandung.

Mardapi, D. (2016). Pengukuran Penilaian dan Evaluasi Pendidikan. Yogyakarta: Parama Publishing.

Massadeh, N. (2020). Internationalization of higher education in Jordan. *Viewpoints Special Edition, Higher Education and*

the Middle East: Building Institutional Partnerships.

McCarthy, J. (2015). Evaluating written, audio and video feedback in higher education summative assessment tasks. *Issues in Educational Research*, *25*(2), 153.

McCarthy, J. (2015). Evaluating written, audio and video feedback in higher education summative assessment tasks. Issues in Educational Research, 25(2), 153-169.

McCarthy, J. (2017). Enhancing feedback in higher education: Students' attitudes towards online and in-class formative assessment feedback models. *Active Learning in Higher Education*, *18*(2), 127–141. https://doi.org/10.1177/1469787417707615

McCutcheon, K., Lohan, M., Traynor, M., & Martin, D. (2015). A systematic review evaluating the impact of online or blended learning vs. face-to-face learning of clinical skills in undergraduate education. *Journal of advanced nursing*, *71*(2), 255-270.

McDonald, J. P. (2002). Dilemmas of planning backwards: Rescuing a good idea. Teachers College Record, 94, 152-169.

Mehmood, Tahir; Hussain, Mubashira Khalid. (2012). Impact of Formative Assessment on Academic Achievement of Secondary School Students, International Journal of Business and Social Science, 3 (17): 101-104.

Ministry of Higher Education and Scientific Research (MOHE, 2020). Brief on Higher Education Sector in Jordan. Retrieved from http://www.mohe.gov.jo/en/pages/BriefMohe2.aspx

Mintz, S. (2020). Reimagining Higher Education Post-Coronavirus. How to make colleges and universities less fragile and higher education more affordable, accessible, equitable, resilient and sustainable. Retrieved from: https://www.insidehighered.com/blogs/higher-edgamma/reimagining-higher-education-post-coronavirus. Accessed 8 September 2020

Mueller, J. (2005). The authentic assessment toolbox: enhancing student learning through online faculty development. Journal of Online Learning and Teaching, 1(1), 1-7.

Ndalichako, Joyce. (2015). Secondary School Teachers' Perceptions of Assessment, International Journal of Information and Education Technology, 5(5): 326-330.

Norcini, J., Anderson, M. B., Bollela, V., Burch, V., Costa, M. J., Duvivier, R., ... & Swanson, D. (2018). 2018 Consensus framework for good assessment. *Medical teacher*, *40*(11), 1102-1109.

Osman, G. (2018). Formal e-Learning in Arab Countries: Challenges and Opportunities. *Learning, Design, and Technology*, 1–26. https://doi.org/10.1007/978-3-319-17727-4_34-1

Pauli, M., & Ferrell, G. (2020). The future of assessment: five principles, five targets for 2025.

Pellegrino, J. W., Chudowsky, N., & Glaser, R. (2001). Knowing what students know: The science and design of educational assessment. National Academy Press, 2102 Constitutions Avenue, NW, Lockbox 285, Washington, DC 20055.

Popham, W. J. (2008). Transformative assessment. VA: ASCD.

Quacquarelli Symonds (2020a). The impact of the coronavirus on global higher education. Retrieved from: http://info.qs.com/rs/335-VIN-535/images/The-Impact-of-the-Coronavirus-on-GlobalHigher-Education.pdf. Accessed 10 September 2020

Reeves, S., Boet, S., Zierler, B., & Kitto, S. (2015). Interprofessional education and practice guide no. 3: evaluating interprofessional education. *Journal of Interprofessional Care*, *29*(4), 305-312.

Regmi, K., & Jones, L. (2020). A systematic review of the factors – enablers and barriers – affecting e-learning in health sciences education, *BMC Medical Education*, 10.1186/s12909-020-02007-6, **20**, 1.

Rhema, A. & Mlliszewska, I. (2010). Towards E-learning in Higher Education in Libya. *Issues in Information Science and Information Technology*, **7**, pp.423-437.

Saye, J. 2013. "Authentic Pedagogy: its presence in social studies classrooms and relationship to student performance on state-mandated tests". *Theory & Research in Social Education* 41: 89-132. doi: 10.1080/00933104.2013.756785

Suharto, PPTK IPA (2015). Teacher Training Materials Implementation Curriculum 2013 Year 2015 Madrasah Aliyah/Senior High School Subjects of Physics, Ministry Education and Culture, Jakarta.

Taras, M. (2005). Assessment - summative and formative - some theoretical reflections. British Journal of Educational Studies, 53(4), 466-478. http://dx.doi.org/10.1111/j.1467-8527.2005.00307.x

Tenison, E., & Touger-Decker, R. (2018). Impact of e-Learning or blended learning versus face-to-face learning in regard to physical examination skills, knowledge, and attitudes among health professions students. *Topics in Clinical Nutrition*, *33*(3), 259-270.

Thorndike, R.M. (2005). Measurement and assessment in psychology and education, (7th ed.). Columbus, OH: Merrill.

Villarroel, V., Bloxham, S., Bruna, D., Bruna, C., & Herrera-Seda, C. (2018). Authentic assessment: Creating a blueprint for course design. *Assessment & Evaluation in Higher Education*, *43*(5), 840-854.

Villarroel, V., Bloxham, S., Bruna, D., Bruna, C., & Herrera-Seda, C. (2018). Authentic assessment: Creating a blueprint for course design. *Assessment & Evaluation in Higher Education*, *43*(5), 840-854.

Vonderwell, S., Liang, X., & Alderman, K. (2007). Asynchronous discussions and assessment in online learning. *Journal of Research on Technology in Education*, *39*(3), 309–328. https://doi.org/1 0.1080/15391523.2007.10782485

Vrazalic, L, Macgregor, J. & Behi, D J. (2009). E-learning Barriers in the United Arab Emirates: preliminary Results from an Empirical Investigation. *IBIMA BUSINESS REVIEW* , **4,** PP. 1-7.

Waseh, S., & Dicker, A. P. (2019). Telemedicine training in undergraduate medical education: mixed-methods review. *JMIR medical education*, *5*(1), e12515.

Westphal, J. (2019). Higher Education in Jordan: Opportunities For Collaboration. Universities UK International. Retrieved from https://www.universitiesuk.ac.uk/policy-and-analysis/reports/Documents/International/2019/he_jordan_2019.pdf

Yeates, P., Cope, N., Hawarden, A., Bradshaw, H., McCray, G., & Homer, M. (2019). Developing a video-based method to compare and adjust examiner effects in fully nested OSCEs. *Medical education*, *53*(3), 250-263.

Zhicheng & Burry-Stock, Judith. (2003). Classroom Assessment Practices and Teachers' Self-Perceived Assessment Skills, *Applied Measurement in Education*, *16*(4): 323–342.

وجهات نظر طلبة الجامعة بطرق تقييم التعلم الممارسة والمتوقعة: تطوير أداة ومسح وطني*

ماجد الخياط* *

ملخص

يعتبر التقويم الفعال أحد الركائز الأساسية اللازمة لتحقيق نتائج التعلم المرجوة، ويمكن أن يكون له تأثيرات جوهرية على أداء الطلبة، وتحفيزهم، واستعدادهم للاستفادة من التجربة التعليمية؛ وعليه فإنه من المهم معرفة ماهية إدراك الطلبة لأساليب التقويم التي يخضعون لها، وما التوقعات لديهم حول ما يجب تطبيقه منها. وعلى الرغم من أهمية البحث في هذا المجال لدعم العملية التعليمية وتطويرها فإنه لا توجد في المنطقة العربية بشكل عام، والأردن بشكل خاص، أي دراسات حتى الآن بحثت في وجهات نظر طلبة الجامعات فيما يتعلق بأساليب التقويم المستخدمة؛ لذلك هدفت الدراسة إلى تطوير وتكييف أداة لقياس طرق التقويم الجامعي، ووصف وجهات نظر الطلبة حول طرق التقويم الأربعة الملاحظة والمتوقعة (الواقعي، التكويني، التشخيصي، النهائي أو الختامي)، إضافة إلى اختبار الاختلافات بين وجهات نظر الطلبة فيما يتعلق بطرق التقويم الملاحظة والمتوقعة، واختبار الاختلافات في وجهات نظر الطلبة فيما يتعلق بأساليب التقويم الملاحظة بناءً على جنسهم، ومجالهم الأكاديمي، ومستواهم الدراسي، تم إجراء مسح مقطعي وصفي وجمع عينة عشوائية طبقية مكونة من(1155) طالبًا وطالبة، تم اتباع سلسلة من الخطوات المنهجية لتطوير وفحص الأداة التي تم تطويرها لقياس طرق التقويم الجامعي، وتم استخدام التحليل العاملي الاستكشافي لفحص بناء الأداة وفقراتها، أشارت النتائج إلى أن طريقة التقويم الأكثر استخدامًا في المرحلة الجامعية هي الطريقة الختامية، تليها الطرق التكوينية، والتشخيصية، والواقعية على التوالي، وكانت درجات التقويم التكوينية والتشخيصية والختامية المتوقعة أعلى بكثير من نظيرتها من الدرجات الملاحظة وبفارق ذو دلالة إحصائية، وقد سجلت الإناث درجات أعلى بشكل ملحوظ في التقويم الواقعي والتكويني والتشخيصي مقارنةً مع الذكور، في حين أظهرت النتائج عدم وجود اختلاف كبير بين الجنسين فيما يتعلق بالتقويم الختامي، كما كانت درجات طلبة الكليات العلمية أعلى بكثير في التقويم الواقعي مقارنةً من نظرائهم في العلوم الإنسانية، بينما كانت النتيجة عكسية في التقويم التكويني، كذلك كانت درجات طلبة الماجستير أعلى بكثير من طلبة البكالوريوس في التقويم الواقعي والتكويني والتشخيصي، بينما لم يتم تُظهر النتائج أي اختلاف فيما يتعلق بالتقويم الختامي. إن نتائج الدراسة مجتمعة تُظهر أن هناك حاجة لتصميم برامج تدريبية متخصصة للهيئة التدريسية في الجامعات في الأردن حول طرق التقويم الحديثة، ومساعدتهم على ربط هذه الأساليب بنتائج التعلم المرجوة وفعالية التدريس، وهنالك أيضًا حاجة إلى دراسات نوعية لبحث معارف واتجاهات الطلبة تجاه طرق التقويم المختلفة، كما أنه من الضروري أن يتم إدماج آراء الطلبة وتوجهاتهم في أي جهود رامية لتطوير عملية التقويم في التعليم العالي

الكلمات الدالة: التقويم الواقعي؛ التقويم التكويني، التقويم التشخيصي، التقويم الختامي.

* أجري هذا البحث بدعم من عمادة البحث العلمي والابتكار بجامعة البلقاء التطبيقية للعام الجامعي 2020/2019

** جامعة البلقاء التطبيقية. تاريخ استلام البحث 2020/7/27، وتاريخ قبوله 2020/12/22.