

**THE DEGREE OF FACULTY'S USE OF AUTHENTIC ASSESSMENT TOOLS AT
AL-QUDS UNIVERSITY AND THEIR RELATION TO THEIR ATTITUDES
TOWARDS THEM**

**Inas Aref Saleh Naser¹, Ibrahim Moh'd Abdel-Rahman Arman¹,
Jamal Subhi Ismail Nafi^{2*}**

¹Department of Education, Faculty of Educational Sciences, P.O. Box: 51000, East
Jerusalem-Abu Dies, Palestine.

²Department of English, Al-Quds University, P.O. Box: 51000, East Jerusalem-Abu Dies,
Palestinian Occupied Territories.

ABSTRACT: *The study aims to investigate and realize the degree of faculty's use of authentic assessment tools at Al-Quds University and their relation to their attitudes towards them. To achieve the purpose of the study, a sample of (99) faculty members at the university was selected in the academic year 2016/2017. Two instruments were developed by the researchers, a questionnaire to measure the degree of faculty's use of authentic assessment tools, and another questionnaire to measure their attitudes towards authentic assessment. And the reliability of the two instruments was reached. The results of the study showed that the use of authentic assessment tools by the faculty at Al-Quds University is moderate, and there were statistically significant differences at ($\alpha \leq 0.05$) in the mean score and in favor of gender and academic rank. And there were no statistically significant differences with regard to experience. The results also showed that there were statistically significant differences in the mean score of the faculty's attitudes towards authentic assessment which reached ($\alpha \leq 0.05$) and was too high, and in favor of academic rank, and there were no significant differences with regard to gender and experience. The study concluded that there is a small positive relationship between the faculty's use of authentic assessment tools and their attitudes towards it. In light of the above results, the study recommends conducting other research about the obstacle that stand on the way of using authentic assessment tools, and carrying out training courses for the faculty about the importance and need of using authentic assessment tools.*

KEYWORDS: Authentic Assessment, Al-Quds University, Attitudes, Faculty

INTRODUCTION

Background of the Study

All kinds of assessment are vital to the learning-teaching process, but not all are effective in the same degree. Authentic assessment is considered to be the most effective kind of assessment that has ever used in academia. Assessment contributes to learning itself; as students are assessed fairly, they will learn better, and their skills will improve. Instructors will also have a clearer idea of students' different levels if they assess them well by using different methods and procedure.

Authentic assessment is defined and understood in different ways by different scholars. Therefore, the researchers thought it is worthy to see how they defined it. Mueller (2005: n.

p.), for example, defined it as “A form of assessment in which students are asked to perform real-world tasks that demonstrate meaningful application of essential knowledge and skills”.

While, Wiggins (1993: 229) is of the view that authentic assessment is "...Engaging and worthy problems or questions of importance, in which students must use knowledge to fashion performances effectively and creatively. The tasks are either replicas of or analogous to the kinds of problems faced by adult citizens and consumers or professionals in the field."

Moreover, Stiggins (1987: 34) thinks that authentic assessment is related to "Performance assessments call upon the examinee to demonstrate specific skills and competencies, that is, to apply the skills and knowledge they have mastered". An authentic assessment usually includes a task for students to perform and a rubric by which their performance on the task will be evaluated.

Keyser and Howell (2008) defined the term authentic as “something real or genuine”. According to them, “When education prepares students for real or genuine tasks, they may encounter in a real-world setting it is considered authentic” (3).

According to Carless, Joughin, Ngar-Fun Liu (2006), assessment affects teaching and learning in a tremendous way. It affects all students and all lecturers in ways that may be positive, benign or negative. They added that “Assessment impacts on what content students focus on, their approaches to learning, and their patterns of study” (2). What they do for their assignments and their preparation for examinations, and how they perceive the results of this assessment, has a profound impact on them both as learners and as individuals. Students follow the cues given to them via assessment and this helps them decide about how to spend their time, so our assessment design needs to be undertaken carefully to increase the positive impact of assessment on the behavior of students at different levels. In short, assessment has such power that it is essential that we handle it so that its learning potential is fully harnessed.

There are different types of assessment, such as Diagnostic Assessment (as Pre-Assessment), Formative Assessment, Summative Assessment, Norm-Referenced Assessment, Criterion-Referenced Assessment and Interim/Benchmark Assessment. All these types of assessment have been practiced in academia since time immemorial.

In this sense it is noteworthy to see what Mueller (2005) says about assessment; he explains the differences between Traditional Assessment and Authentic Assessment. By "traditional assessment" (TA) he referred to the “forced-choice measures of multiple-choice tests, fill-in-the-blanks, true-false, matching and the like that have been and remain so common in education” (8). Students typically select an answer or recall information to complete the assessment. These tests may be standardized or teacher-created. They may be administered locally or statewide, or internationally.

Behind traditional and authentic assessments is a belief that the primary mission of schools is to help develop productive citizens. From this common beginning, the two perspectives on assessment diverge.

In contrast, authentic assessment (AA) springs from the following reasoning and practice: First, a school's main mission is to develop productive citizens. Second, to be a productive citizen, an individual must be capable of performing meaningful tasks in the real world. Third, schools must help students become proficient at performing the tasks they will encounter when they

graduate. Finally, to determine if it is successful, the school must then ask students to perform meaningful tasks that replicate real world challenges to see if students are capable of doing so.

Thus, in AA, assessment drives the curriculum. That is, teachers first determine the tasks that students will perform to show their mastery, and then a curriculum is developed and modified which may enable students to perform those tasks very well, and that would include the acquisition of essential knowledge and skills.

But a teacher does not have to choose between AA and TA simply because both complement each other. He/she is likely to blend both methods so as to achieve goal and desired ends.

The aim of this study is to investigate the degree faculty's use of authentic assessment tools at Al-Quds University and their attitudes towards them. To achieve the purpose of the study, two questionnaires were used, one to measure the degree of faculty's use of authentic assessment tools, and another to measure faculty's attitudes towards using these tools.

Statement of the Problem

Assessment is no longer limited to identifying the amount of what the student has achieved in the course material, but rather to the knowledge of all the characteristics of the student's personality and directing his learning towards the development of higher order thinking skills and employing knowledge to solve various life problems. This in turn has made authentic assessment take its place in the educational process by providing the appropriate measuring tools for authentic learning situations. And since the university is an institution to build an integrated personality of the learner, it was important to study the methods and tools used by the faculty to assess students' performance, which prompted the researchers to study the degree of the faculty's use of authentic assessment tools and their relationship and attitudes towards them at Al-Quds University.

Questions of the Study

The study tries to answer the following questions:

1. What is the degree of the faculty's use of authentic assessment tools at Al-Quds University?
2. Are there any statistically significant differences in the degree of faculty's use of authentic assessment tools at Al-Quds University due to gender, experience and qualification?
3. What are the attitudes of the faculty members at Al-Quds University towards authentic assessment?
4. Are there any statistically significant differences in the attitudes of the faculty members at Al-Quds University due to gender, experience and qualification?
5. What is the relationship between the degree of the faculty's use of authentic assessment tools and their attitudes towards it?

Hypotheses of the Study

The study tries to test the following hypotheses:

1. There are no statistically significant differences at the level of significance ($\alpha \leq 0.05$) in the arithmetical averages of the degree of faculty's use of authentic assessment tools due to gender.

2. There are no statistically significant differences at the level of significance ($\alpha \leq 0.05$) in the arithmetical averages of the degree of faculty's use of authentic assessment tools due to experience.
3. There are no statistically significant differences at the level of significance ($\alpha \leq 0.05$) in the arithmetical averages of the degree of faculty's use of authentic assessment tools due to qualification.
4. There are no statistically significant differences at the level of significance ($\alpha \leq 0.05$) in the arithmetical averages of the faculty's attitudes towards authentic assessment due to gender.
5. There are no statistically significant differences at the level of significance ($\alpha \leq 0.05$) in the arithmetical averages of the faculty's attitudes towards authentic assessment due to experience.
6. There are no statistically significant differences at the level of significance ($\alpha \leq 0.05$) in the arithmetical averages of the faculty's attitudes towards authentic assessment due to qualification.
7. There are no statistically significant differences at the level of significance ($\alpha \leq 0.05$) between the degree of the faculty's use of authentic assessment tools and their attitudes towards them.

Objectives of the Study

The study tries to achieve the following goals:

1. Determining the degree of faculty's use of authentic assessment tools at Al-Quds University.
2. Knowing the effect of each of the variables (gender, experience and qualification) in determining the degree of faculty's use of authentic assessment tools.
3. Knowing the attitudes of faculty members at Al-Quds University towards authentic assessment.
4. Knowing the impact of each of the variables (gender, experience and qualification) on the attitudes of faculty members at Al-Quds University towards authentic assessment.
5. Examining the relationship between the degree of faculty's use of authentic assessment tools and their attitudes towards them.

Significance of the Study

The study is considered significant for the following reasons:

1. The nature of the subject that is being dealt with, as it deals with the authentic assessment and its tools.
2. This study may open up considerable scope for further research and studies on the use of authentic assessment tools, and on the constraints and problems teachers face in their use.

3. This study may contribute to helping teachers to rethink their methods of evaluating students and to change them and use modern assessment tools.
4. This study may help the university to change and develop the current assessment tools, and use other tools focusing on the authentic skills and performance of the student.

Limitations of the Study

The study is limited to a sample of faculty members at Al-Quds University in the academic year 2016/2017. The degree to which faculty members at Al-Quds University use authentic assessment tools and their attitudes towards them is determined by the extent to which they respond to the statements in the two questionnaires.

Related Studies

Croft, Mary and Vera (2016) conducted a study to investigate the perfect storm of education reform: high-stakes testing and teacher evaluation. The article examined seemingly disconnected education reform policies and posits that their unprecedented alignment is eroding the bedrock of public education. Using Georgia as an example, the authors demonstrate how neoliberal efforts to reform education occur through three systematic and interconnected fronts: political climate change, the testing industrial complex, and a mesoscale evaluation system. The authors challenge assertions that those reforms increase academic achievement and global competitiveness. Instead, the orchestrated alignment is being experienced as an assault on the supposed beneficiaries (i.e., public education and teacher education). These conceptual weather fronts can serve as a means to analyze stated intentions versus outcomes of education policy. The authors concluded with grassroots responses by students, teachers, and others to the destructive elements of reform. One left office shrouded in a legacy of disdain for teachers; the other left office with educators realizing that public education had been significantly weakened

Zu'bi (2013) did a study that aimed at investigating the degree that mathematical teacher's knowledge and utilization authentic assessment at the high primary grades. To achieve the purpose of this study, a special questionnaire was developed by the researcher consisted of five parts, the validity and reliability of the questionnaire were checked also. The questionnaire was applied on a sample of (91) teachers in Irbid Directorate of Education, First Department, observation card was also applied on a sample of (27) teachers in Irbid Directorate of Education, First Department. The results showed that the

practicing degree of authentic assessment strategy by mathematics teachers was moderate except for the use of observation strategy which shows high level. There were no significant differences identified by respondents due to gender, academic degree and experience. The most prominent constraint to the use of authentic assessment strategies that: it is time consuming. The study included many recommendations. The most important one was the necessity of providing training programs for teachers in the field of modern assessment strategies in order to improve the ability of teachers in this field.

Alavi (2014) investigated standardized testing of the non-standardized Arabic-speaking ELLs. The rapidly growing population of English Language Learners (ELLs) has brought new challenges to schools throughout the United States. Research has also demonstrated a disparity in achievement between ELLs and the general student population in association with the increasing accountability demands of No Child Left Behind (NCLB, 2001). The investigation

conducted in a charter school serving students of predominately Arabic descent showed not only a misalignment between students and staff views of testing, it also demonstrated that despite all our best efforts, the school continued to yield to the pressures of standardized assessments. Teachers, administrators and students from a metro area in the Midwest participated via archival data, staff questionnaires and staff and student one-on-one interviews. The purpose of this research was to explore standardized testing practices on ELLs, specifically those of Arabic speaking backgrounds, and to analyze the role standardized tests play on the instructional time needed for Arabic-speaking ELLs to acquire CALP. Non-native English students take time, which can range from five to seven years, to become proficient in the academic language. In the current study it was found that teachers forfeited countless hours of much needed instructional time, in order to accommodate preparing for, and administering of, these assessments. Moreover, while attempts were made to minimize emphasis on testing, the school succumbed to the ongoing demands of external stakeholders. Consequently, Arabic-speaking ELLs were not receiving the instructional time necessary for them to acquire the English language proficiency which was equivalent to that of their English speaking counterparts.

Hago and Ali (2014) conducted a study that is restricted to assess writing performance in the Sudanese female students at secondary schools at Omdurman locality, Aluola secondary school for girls, in the academic year 2013 – 2014. It aims at investigating the English syntactic structures experienced by Sudanese Students at secondary schools. The researcher used the analytical descriptive method in this study and a test as a tool for collecting data. The sample of the study was about ninety nine students at secondary schools in the academic year (2013-2014). After the analysis of the types of errors made by the subjects, the study has come out with many findings. The major problem behind the students' errors is the mother tongue interference. The Sudanese learners of English in general seem not to have an adequate proficiency in understanding the meaning and semantics when they express themselves in English syntactic structures. Students need a supplementary method in order to express themselves accurately. Finally the researcher has recommended certain areas such as: Teachers and students should be aware of the importance of writing in relation to other skills. Activating English literature lessons and providing a library for extra activities. Students need enough time to practice writing in the class room because the time allotted for teaching English is not matched to the content of the syllabus designed. Students should be prepared to use the language for a variety of purposes beyond the classroom.

According to Rita (2008), assessment has to be seen as an interconnected part of teaching and learning. The conception of assessment is one that focuses on describing student learning, identifying where each student is in his or her personal learning progression, diagnosing any difficulties students may be having in their learning, and providing direction to the teacher and the student in the steps to be taken to enhance learning. This focus on the use of assessment to support learning, rather than to document achievement, has come to be referred to as "assessment for learning" (3). To bring the concept of assessment for learning to fruition in the classroom, assessment activities have to be designed and conducted with the purpose of learning in mind. If teachers are to embrace this new philosophy of assessment, they need to understand how assessment for learning works.

Afaneh (2011) did a study with an objective to determine recent trends in the evaluation. It also aimed to identify the reality Arabic language teachers use, in the preparatory stage in UNRWA schools in the Gaza Strip, of alternative evaluation, and determine whether different teachers

use the methods of evaluation depending on the gender variable. To achieve these goals the researcher followed the descriptive approach, using two tools, namely: the questionnaire and note card. The study sample included (60) teachers from the east and the West of the Gaza Strip. It included also sample of (24) managers and Supervisors. The researcher noted the reality of the use of Arabic language teachers, in UNRWA schools in Gaza City, of the alternative methods of evaluation. A questionnaire was distributed among managers and supervisors for the same purpose. To address the data statistically, the researcher used the Statistical Packages for Social Sciences (SPSS) program. The researcher used the frequency of occurring, percentages, the arithmetic means, standard deviation, the relative weight, and the T. test, and Pearson correlation coefficient, and coefficient of stability. The researcher concluded that the total marks for the use of teachers (male) for the modern evaluation methods amounted to (45.1%), while the seventh dimension of the observation card (performance evaluation tests written) ranked first relative by a weight of (70.5%), which is the highest among the relative weights, followed by a fifth dimension (evaluation of performance based on the observation) and its relative weight was the (56.7%), followed by the third dimension (self-evaluation) and it weighs (58.5%), followed by the first dimension (evaluation based on performance) and weighs (55.8%), followed by the fourth dimension (peers; evaluation) and its relative weight was (46.7%), followed by the eighth dimension (reformation of performance by maps concepts) and it weighs (23.2%), followed by the sixth dimension (evaluation of performance by interviews) and weighs (21.1%), followed by the second dimension (business files), its relative weight was (20%). The study recommended that review of the current evaluation practices that rely on traditional tests is no longer acceptable for the teachers to continue to understand that the evaluation is a synonym for the exams; and the role of the school continues to be limited in the scope of preparing the students for the tests rather than the understanding, Developing curricula, organizing consecutive training evaluation seminars for female and male of Arabic language teachers at alternative methods of evaluation to improve the quality of their performance when they assess students in various branches of the Arabic language, etc.

Boud and Falchikov (2005) presented a paper at a conference and believed that an important rationale for higher education is that it equips students for learning beyond the point of graduation. The paper considers the role that assessment plays in this. It suggests the need to take a new perspective on assessment: assessment to promote learning throughout life. The paper focuses on ideas that can be used to contribute to the construction of assessment practices and on wider implications for course design. It concludes by exploring barriers to acceptance of this perspective and how they might be addressed.

Donovan, Larson, Beth, Stechschulte and Taft (2002) conducted an action research that demonstrated that alternative assessments are a more comprehensive way of exhibiting student achievement. Alternative Assessment is a non-standardized method of assessment. Among other types of evaluations, this definition includes such things as portfolios, checklists, rubrics, surveys, student involved assessments, as well as reflections. The selected student population consisted of fifth, seventh and eighth grade classes. Alternative Assessment skills were assessed and documented to determine an increase in student involvement and enthusiasm toward work. Analysis of standardized testing shows a limited profile of actual student achievement. More importantly, they fell short in measuring student ability in active skills such as writing, speaking, acting, drawing, constructing, repairing, and other skills that are required of students over the course of their schooling. Standardized tests generally focused on the final answer instead of the learning processes involved in getting to that final answer. A review of

solution strategies suggested by our research, included the use of portfolios, questionnaires, rubrics, various student self assessments, as well as student and teacher reflections. These assessment skills were taught at the beginning and reinforced throughout the intervention process. They concluded that post intervention data indicated using various types of evaluation is a complete and, more importantly, authentic assessment of student achievement.

Chan and Gurnam (2010) have noted that there is a mismatch between curriculum content and assessment practices in higher education. At the moment, the focus is still on the assessment of learning and not much on assessment for learning. Therefore, this study was conducted to examine the implementation of authentic assessment in higher education in Malaysia. Approach: The study employed a qualitative research method and involved the use of instruments such as interviews, document analysis and unobtrusive classroom observations to collect the relevant data in the classrooms of a three-credit hour course from a Masters Degree program. The total population consisted of 2 lecturers and 20 students from the two selected intact TSL 752 classes. Results: In this study, researchers identified different types of authentic assessment with the suitability to certain pedagogical strategies, their effects on students' learning and the appropriate procedures of conducting authentic assessments. The findings indicated that alternative and authentic assessment have more acceptance from students and should therefore be viewed as an alternative to traditional standardized assessment. The study concluded that assessment strategies should be closely related to teaching and learning. Assessing authentic performances should become integral parts of the instructional cycle and feedback provided by the lecturer and peers should be formative in order to help the students assess their strengths and weaknesses, identifying areas of needed growth and mobilizing current capacity.

Fox, White and Kidd (2011) did a qualitative study that examined the portfolio reflections of 51 teachers enrolled in an advanced master's degree program whose learning outcomes are aligned with the core propositions of the National Board for Professional Teaching Standards. Using the four levels of reflection and a fifth level that emerged from data analysis, we examined the levels of development and change in teachers' reflections across the program. This analysis helped the researchers to determine to what degree teachers developed a reflection-based inquiry stance in their classrooms during the program. The study contributes new findings to the body of literature on the role and function of portfolios in teacher education programs seeking to document teachers' understanding and application of specific program goals and professional standards as a result of professional development coursework. As an authentic assessment tool, portfolio data (i.e., evidence such as the reflections contained therein) can provide an important lens for capturing teachers' approach to teaching and learning and provides insight into the complexity of professional development for practicing teachers.

Gijbels, Watering, Dochy and Bossche (2005) carried out a study to get more insight in the effects of written assessment tasks integrated in a problem-based learning environment. Both the influence on students' performances and students' perceptions were investigated. Students' final exam results were used to find out whether students who make the assessment tasks do better than students who do not. Answers from questionnaires and semi-structured interviews were used to discover the most important concerns in students' and teachers' perceptions of the assessment tasks. The results indicate that making the assessment tasks had positive influence on the students' overall performance. From the questionnaires and interviews it appears that both the students and the teachers see the benefits of the assessment tasks. It is concluded that

small steps in the change of the assessment system can result in relatively big changes in students' learning and results.

Keyser and Howell (2008) conducted a review to briefly summarize the state of authentic assessment as it evolves theoretically and then integrates itself into educational practice. Some of the questions it seeks to answer from the literature include: how did it begin; what were the circumstances surrounding its inception; how does it relate to performance and alternative assessment; what are obstacles to wider adoption; what impact is technology having in its emergence; and what should happen next? The article identifies authors and cites studies that have informed the discussion on authentic assessment. It is evident from the review that authentic assessment is an important educational concept that at the time of this writing is in a fledgling state of definition and integration. They concluded that Authenticity is the element of every successful assessment that resembles a real-world skill or activity and aligns itself with a learning outcome. Since educational reforms in the late 1980s, educators and researchers alike have shown increasing interest in defining and applying authentic assessment. These emerging theoretical and applied models promise to inform future research and best practices.

Lanting (2000) carried out a qualitative research to examine how four primary teachers used a district literacy performance assessment. Data were collected through observations, interviews, and documents. Grounded theory and NUD*IST software were used for text analysis and theory building. Findings show that a theory-grounded teacher-empowered K-2 performance assessment program accompanied by the school district's low level of interference could interact very well with teachers' high ethical standards on assessments. When the in-service training was voluntary, teachers did not spontaneously practice comprehended portfolio use or students self-assessment, but rather relied on observation and interview. Rubrics seemed to work as conceptual frameworks for data collection and evaluation, and teachers usually grounded their evaluations on evidence. Dimensional scoring and flexible marking across proficiency levels were implemented, and teachers appeared to focus on student strengths. The information obtained from the literacy assessment was criterion referenced and individualized. Teachers did not use normative language when commenting on student performance. Assessment results were generally used to keep track of student performance and to provide remedial teaching, but there seemed to be a gap between assessment results and corresponding pedagogical strategies. It is recommended that performance assessment programs be accompanied by teacher in-service education on repertoires of instructional strategies, but it is not clear whether teacher involvement in rubric development and peer discussions would make assessment results more instructionally useful.

Marzano (2002) did a set of studies compared the findings from generalizability (G) studies and alternative decision (D) studies for 4 approaches to scoring classroom assessments. In the 1st experiment, the unconstrained point method was compared to the generic rubric method. Under the unconstrained point method, raters assigned total possible points to items and then scored the responses of 10 students to those items to compute an overall percentage score for each student. Under the generic rubric method, raters were given a generic rubric that they applied to student responses for all items as a set. In the 2nd experiment, the constrained point method was compared to the topic-specific rubric method. Under the constrained point method, raters scored the responses of 10 students to items whose point values had been previously established. In the topic-specific rubric condition, raters used a rubric that was specific to the topic of the assessment. In terms of less Rater \times Person variability and higher G and D coefficients, the topic-specific rubric method was found superior to the constrained point

method that was found superior to the generic rubric method that was found superior to the unconstrained point method.

Oladele (2011) examined knowledge and use of authentic assessment techniques (a performance based assessment requiring learners to utilize their knowledge in a meaningful context) among lecturers in Botswana College of Agriculture. A simple random sampling technique was used to select 40 lecturers from 96. The results show that in terms of use, 47.5% of lecturers use authentic assessment while 52.3% do not. Lecturers were more knowledgeable on items such as authentic assessment requires students to apply skills and abilities as they would in real life (100%); and authentic assessment involves direct examination of student's ability to use knowledge to perform a task that is encountered in real life (100%). The most prominent constraint to the use of authentic assessment is that it is time consuming (90%). Significant determinants were teaching experience ($t = 3.61$), educational level ($t = 4.36$), holding administrative positions ($t = -4.71$) and knowledge of authentic assessments ($t = -3.90$). It is important therefore that the use of authentic assessment is popularized among lecturers.

Methods and Procedures

Methods

The associative descriptive approach was applied in this research, in accordance with its objectives in determining the degree of faculty's use of authentic assessment tools and their attitudes towards them and the relationship between them by applying the tools to the sample members to reach the results that answer the research questions and test its hypotheses.

Population of the Study

The population of the study included all faculty members (8600) at Al-Quds University in the academic year 2016/2017.

Sample of the Study

The sample of the study consisted of (99) faculty members at Al-Quds University, randomly selected by (25%) of the study population. Table (1) shows the characteristics of the demographic sample.

Table (1): Characteristics of the demographic sample

Variable	Variable levels	Number	Percentage
Gender	Male	69	70
	Female	30	30
Experience	1-5 years	11	11,1
	5-10 years	11	11,1
	10-15 years	31	31,3
	15 years and above	46	46,5
Qualification	Lecturer	11	11,1
	Assistant Professor	41	41,4
	Associate Professor	29	29,3
	Professor	18	18,2
Total		99	100

Variables of the Study

Independent variable

- Gender (Male, Female)
- Experience (1-5 years, 5-10 years, 10-15 years, 15 years and above)
- Qualification (Lecturer, Assistant Professor, Associate Professor, Professor).

Dependent variable

- The degree of faculty's use of authentic assessment tools.
- The attitudes of the faculty towards authentic assessment.

Instruments of the Study

The researchers prepared the instruments of the study (Questionnaire to measure the faculty's use of authentic assessment tools, and another questionnaire to measure their attitudes towards authentic assessment) by consulting other studies such as Anzi's (2016), Hamza and Soman's (2012), Abu Khalifa et al. (2011) until the instruments were finally developed in their final forms. The reliability of the instruments was achieved by exposing them to a jury of experienced professors of the same field, their recommendations and suggestions were taken into consideration. Stability coefficient (Cronbach Alpha) was also calculated to determine the degree of faculty's use of authentic assessment tools which was (0.80), while the coefficient for the attitudes questionnaire was (0.86).

Statistical Analysis

Statistical analysis of the data was done using numbers, percentages, computational averages, standard deviations, t-test, one way analysis of variance, Cronbach Alpha and Pearson correlation coefficient using SPSS. The following debug key was used:

low $2.33 \geq$ Arithmetic Average

3.66 Moderate \geq Arithmetic Average

3.66 High \geq Arithmetic Average

RESULTS AND DISCUSSION

What is the degree of faculty's use of authentic assessment tools at Al-Quds University?

To answer this question, the researchers calculated the mean and standard deviation, with a mean of (3.43) and a standard deviation of (0.77).

The researchers believe that the reason for this result is that there are some evaluation tools used by the faculty members at the university to a large extent, such as paper tests, pen and project, while there are very few tools used such as the portfolio and the story file, making the overall score medium.

This result is not consistent with Croft et al. (2016) who examined the education reform by testing and evaluation.

2. Are there statistically significant differences in the degree of faculty's use of authentic assessment tools at Al-Quds University due to gender, experience and qualification?

To answer this question, it was turned into the following hypothesis: There were no statistically significant differences at the level of significance ($\alpha \leq 0.05$) in the arithmetical averages of the degree of faculty's use of authentic assessment tools due to gender.

To examine the hypothesis, the researchers used t-independent sample analysis, as shown in Table 2.

Table (2): Results of the t- independent sample analysis of the respondents' responses to their use of authentic assessment tools due to gender.

Variable	Number	Mean arithmetic average	Standard Deviation	Degree of freedom	Calculated-t value	Level of calculated significance
Male	69	3.31	0.75	97	2.42	0.017
Female	30	3.71	0.77			

The above table shows that the calculated level of significance and its value is (0.017) which is less than the statistical level of significance ($\alpha \leq 0.05$). Therefore, the null hypothesis is rejected and the alternative hypothesis is accepted with statistically significant differences and in favor of females.

The researchers attribute this finding to the fact that females are fewer than males in the university. In addition, the majority of females in the humanities colleges have used different assessment tools that focus on the integrated personality of the learner.

This result is consistent with Boud and Falchikov (2005) who conducted a study to evaluate the importance and use of assessment tools by genders in higher education.

Results of the second null hypothesis: There were no statistically significant differences at the level of significance ($\alpha \leq 0.05$) in the arithmetical averages of the degree of the faculty's use of authentic assessment tools due to experience.

To test the hypothesis, the arithmetical average and the standard deviation were calculated for the responses of the study sample members to the degree of their use of authentic assessment tools as shown in Table (3).

Table (3): Arithmetic average, standard deviation and the preparation of the sample members due to experience.

Level	Number	Arithmetic value	Standard Deviation
1-5 years	11	3.57	0.56
5-10 years	11	3.56	0.32
10-15 years	31	3.37	0.69
15 years and above	46	3.41	0.97
Total	99	3.43	0.77

It is noted from Table (3) that there are apparent differences in the sample, and to know the source of the differences, the analysis of variance (one way ANOVA) was used, as shown in Table 4.

Table (4): Results of the analysis of the single variance of the degree of faculty's use of authentic assessment tools by the sample members due to experience.

Source of Contrast	Total squares	Degrees of freedom	Average squares	Calculated F value	Level of significance
Between groups	0.55	3	0.18	0.30	0.83
Within groups	58.11	95	0.61		
Total	58.66	98			

The statistical significance level is (0.83) which is greater than the statistical significance level ($\alpha \leq 0.05$). Therefore, the null hypothesis is accepted because there are no differences between the sample members due to experience.

The researchers attribute this finding to the fact that all faculty members work in the same educational environment and are subject to the same experiences, in addition to the modern subject of authentic assessment, which made the teacher's experience a variable that did not affect the degree of faculty's use of authentic assessment tools.

This result is consistent with Zu'bi's (2013) study which investigated the degree that mathematical teacher's knowledge and utilization of authentic assessment at the high primary grades.

The third null hypothesis: There are no statistically significant differences at the level of significance ($\alpha \leq 0.05$) in the arithmetical averages of the degree of faculty's use of authentic assessment tools due to qualification.

To test the hypothesis, the arithmetical averages and the standard deviations were calculated for the responses of the sample to the extent of their use of authentic assessment tools due to qualification. Table (5) shows that.

Table (5): Arithmetic averages, standard deviations, and the preparation of the degree of the faculty's use of authentic assessment tools by the sample members due to qualification.

Level	Number	Arithmetic Average	Standard Deviation
Lecturer	11	3.57	0.56
Assistant Professor	41	3.65	0.67
Associate Professor	29	3.57	0.54
Professor	18	2.61	0.91
Total	99	3.43	0.77

Table (5) shows that there are apparent differences, and to know the source of these differences, the analysis of the mono-variance was used, as shown in Table (6).

Table (6): Results of the analysis of the variance for the degree of faculty's use authentic assessment tools due to qualification.

Source of Contrast	Total Squares	Degrees of Freedom	Average squares	Calculated F value	Level of significance
Between groups	14.76	3	4.92	10.65	0.001
Within groups	43.89	95	0.46		
Total	58.66	98			

Table (6) shows that the calculated level of significance and its value is (0.001) which is lower than the level of statistical significance ($\alpha \leq 0.05$). Therefore, the null hypothesis is rejected, and the alternative one is accepted. There are statistically significant differences according to the variable of the degree and to know these differences, Post-hoc (LSD) is used as shown in Table (7).

Table (7): Results of analysis of the differences of dimension due to qualification (LCD).

Level of variables	Difference in averages
Lecturer Assistant Professor	-0.08
Associate Professor	-0.01
Professor	* 0.95
Assistant Professor Lecturer	0.08
Associate Professor	0.07
Professor	* 1.03
Associate Professor Lecturer	0.01
Assistant Professor	-0.07
Professor	* 0.96
Professor Lecturer	* -0.95
Assistant Professor	* -1.03
Associate Professor	* -0.96

* Statistical Significance

It is shown in Table (7) that comparing the lecturer with the professor, the results is in favor of the professor, comparing the assistant professor with the professor, the result is in favor of the assistant professor, and comparing the professor with the associate professor, it is in favor of the associate professor.

The reason for this finding is that the professor is better informed than the lecturer of his activity and his research experience. The assistant professor and the associate professor are often more familiar with recent subjects than the professor by their achievements in research for the purpose of promotion.

This result is consistent with Alavi (2014) and Afaneh (2011) who studied the recent trends in the evaluation and the various tools used by different school and college teachers to assess their students' academic performance.

3. What are the attitudes of faculty members at Al-Quds University towards authentic assessment?

To answer this question, the researchers calculated the arithmetic average and standard deviation; the total arithmetic average is (4.12) and the standard deviation is (0.54) which were high.

The researchers explain this finding that faculty members at Al-Quds University have positive attitudes towards the use of authentic assessment tools in teaching, but there is insufficient knowledge of how these tools are constructed and used, in addition to the obstacles to using them.

This result is not consistent with Rita's (2008) study which assessed the assessment tools and their relation to teaching and learning.

4. Are there any statistically significant differences in the attitudes of the faculty members at Al-Quds University due to gender, experience and qualification?

To answer this question, it was transformed into the following null hypothesis:

There are no statistically significant differences at the level of significance ($\alpha \leq 0.05$) in the arithmetical averages of the attitudes of faculty members at Al-Quds University towards authentic assessment due to gender.

To test the hypothesis, the researchers used independent sample tests as shown in Table (8).

Table (8): Results of the analysis of the independent samples of the trends of the sample members towards authentic assessment due to gender.

Variable	Number	Arithmetic Average	Standard Deviation	Degree of Freedom	Calculated t-value	Level of significance
Male	69	4.07	0.52	97	1.40	0.16
Female	30	4.23	0.57			

Table (8) shows that the level of significance and its value (0.16) which is greater than the level of statistical significance, and therefore the null hypothesis is accepted. There are no differences between the faculty members in their attitudes towards authentic assessment due to gender.

The researchers attribute this finding to the fact that faculty members, regardless whether they are male or female, live in the same environment and are exposed to the same experiences and, therefore have the same attitudes towards authentic assessment.

This result is consistent with Oladele (2011) who examined the knowledge and use of authentic assessment techniques by both genders.

The Fifth Null Hypothesis:

There were no statistically significant differences at the level of significance ($\alpha \leq 0.05$) in the arithmetical averages of the attitudes of the faculty members at Al-Quds University towards authentic assessment due to experience.

To test the hypothesis, arithmetical average and standard deviation were calculated according to experience, as shown in table (9).

Table (9): Arithmetic average, standard deviation and the preparation of the trends of the sample members towards authentic assessment due to experience.

Level	Number	Arithmetic value	Standard Deviation
1-5 years	11	3.91	0.66
5-10 years	11	4.04	0.14
10-15 years	31	4.17	0.57
15 years and above	46	4.15	0.53
Total	99	4.12	0.54

Table (9) shows that there are apparent differences between the sample members, and to identify the source of the differences, analysis of mono – variance was used, as shown in table (10).

Table (10): Results of the analysis of the single variance of the trends of the sample towards the use of authentic assessment due to experience.

Source of Variance	Average Squares	Degree of Freedom	Average squares	Calculated F value	Level of significance
Between groups	0.69	3	0.23	0.78	0.51
Within groups	27.85	95	0.29		
Total	28.54	99			

It is clear from Table (8) that the level of significance calculated and its value (0.51) which is greater than the level of statistical significance and, therefore, the null hypothesis is accepted. There are no differences in the attitudes of the faculty towards authentic assessment due to experience.

The researchers attribute this finding to the fact that faculty members at the university, depending on the number of their years of experience, are exposed to the same experiences and undergo the same training, making the experience a variable that does not affect their attitudes towards authentic assessment.

This result is consistent with Lanting (2000) and Marzano (2002) who carried out action researches to examine the attitudes of faculty members towards assessment.

There are no statistically significant differences at the level of significance ($\alpha \leq 0.05$) in the arithmetic average of the degree of faculty's attitudes towards authentic assessment due to qualification.

To test the hypothesis, arithmetic average and standard deviation were calculated due to qualification, as shown in Table (11).

Table (11): Arithmetic average, standard deviation and the preparation of the trends of the sample members towards authentic assessment due to qualification.

Level	Number	Arithmetic Average	Standard Deviation
Lecturer	11	3.91	0.66
Assistant Professor	41	4.18	0.49
Associate Professor	29	4.27	0.45
Professor	18	3.85	0.59
Total	99	4.12	0.54

Table (11) shows that there are apparent differences between the sample and, to identify the source of the differences, Analysis of mono – variance was used, as shown in table (12).

Table (12): Results of the analysis of the single variance of the trends of the sample towards the use of authentic assessment due to qualification.

Source of Variance	Average Squares	Degrees of Freedom	Average squares	Calculated F value	Level of significance
Between groups	2.68	3	0.89	3.28	0.024
Within groups	25.86	95	0.27		
Total	28.53	98			

It is clear from Table (12) that the level of significance calculated and the value of (0.024) is less than the level of statistical significance and, therefore, the null hypothesis was rejected. To detect for whose benefit these differences are for, the advanced statistical analysis Post-hoc (LSD) was used, as shown in Table (13).

Table (13): Results of the LSD analysis of the distance differences due to qualification.

Level of variables		Difference in averages	
Lecturer	Assistant Professor	0.27-	0.06
	Associate Professor	0.36-	
	Professor		
Assistant Professor	Lecturer	0.27	* 0.33
	Associate Professor	0.09-	
	Professor		
Associate Professor	Lecturer	0.36	* 0.42
	Assistant Professor	0.09	
	Professor		
Professor	Lecturer	0.06	- 0.42 *
	Assistant Professor	* 0.33-	
	Associate Professor		

* Statistical significance

Table (13) shows the comparison of the assistant professor with the Professor in favor of the assistant professor, and comparing the associate professor with the professor in favor of the professor.

This finding can be attributed to the fact that the assistant professor is usually more motivated to the teaching process and change than the professor, and with regard to the associate professor may be due to his concern to conduct research for promotion, while the professor is more spontaneous and stable in the teaching process.

This result is consistent with Keyser and Howell (2008) who studied the state of authentic assessment and the way it is practiced by teachers at different levels in education.

What is the relationship between the degree of the faculty's use of authentic assessment tools and their attitudes towards them?

To answer this question, it was turned into the following hypothesis:

The seventh null hypothesis:

There was no statistically significant relationship at the level of significance ($\alpha \leq 0.05$) between the degree of the faculty's use of authentic assessment and their attitudes towards them.

Thus, the Pearson correlation coefficient (R) was calculated between the degree of the faculty's use of authentic assessment and their attitudes towards it. The Pearson correlation coefficient (0.1) indicates a weak positive relationship.

The researchers attribute this finding to the fact that the use of authentic assessment tools in teaching is indicative of positive attitudes towards them and weak because there are positive attitudes of teachers towards using authentic teaching, but do not have the knowledge of how to build tools and use strategies.

This result is consistent with Afaneh's (2011) study which investigated the employment of authentic assessment tools by teachers and their attitudes towards them.

CONCLUSION AND RECOMMENDATIONS

The study concluded that:

1. The degree of faculty's use of authentic assessment tools is medium, while the degree of their attitudes towards it is significant.
2. There are statistically significant differences in the degree of faculty's use of authentic assessment tools due to gender and qualification, while there are no differences due to experience.
3. There are statistically significant differences in the attitudes of faculty members towards the use of authentic assessment due to qualification, while there are no differences attributed to gender and experience.
4. There is a weak positive relationship between the degree of faculty's use of authentic assessment tools and their attitudes towards them.

RECOMMENDATIONS

In light of the above results, the researchers recommend:

1. Conducting training workshops for faculty members at the university on the tools of authentic assessment and how to prepare and employ them.
2. Conducting other studies on authentic assessment, focusing on the obstacles and problems that prevent their use.

REFERENCES

Abu Khalifa, I., Khader, G., Isha, I., & Hammash, H. (2011). The degree of the employment of the teachers of the first basic seminar for authentic assessment

- tools and strategies in the schools of Amman Governorate - Jordan from the point of view of the teachers themselves. *Journal of Studies in Education*, 38: 984-1002. Retrieved from <https://search.mandumah.com/record>
- Afaneh, M. A. (2011). *The Use of Basic Assessment Tools by Instructors in the Unrwa Secondary Schools in Gaza Strip in Light of Modern Trends*. (Published MSc. Thesis). The Islamic University, Gaza, Palestine. Pp. 1-225. Retrieved from http://library.iugaza.edu.ps/browse_thesis.aspx?college=4&department=402
- Alavi, A. F. (2014). *Standardized Testing of the Non-Standardized Arabic-Speaking ELLs: A Misalignment of Perceptions*. University of Michigan-Dearborn (Published doctoral thesis).
- Anzi, F. (2016). The extent to which social studies teachers use authentic assessment strategies and tools in Hafr al-Batin Governorate from the perspective of the teachers themselves. *Journal of the Faculty of Education in Educational Sciences*, 40(1): 13-64. Retrieved from <https://search.mandumah.com/record/773300>
- Berry, R. (2008). *Assessment for Learning (Chapter: Basic Concepts of Assessment)*. Hong Kong: Hong Kong University Press. Pp.1-21. Retrieved from <http://www.jstor.org/stable/j.ctt1xcs68.6>
- Boud, D., & Falchikov, N. (July 2005). Redesigning assessment for learning beyond higher education. *Higher Education Research and Development*, 28, 34-41. Retrieved from <https://www.researchgate.net/publication/228337704>
- Chan, Y. F., & Gurnam, K. S. (2010). Authentic Assessment and Pedagogical Strategies in Higher Education. *Journal of Social Sciences* 6(2): 153-161, 2010. Retrieved from <http://thescipub.com/PDF/jssp.2010.153.161.pdf>
- Croft, S. J., Roberts, A. M., & Stenhouse, V. L. (2016). The perfect storm of education reform: high-stakes testing and teacher evaluation. *Social Justice*, 42(1): 70-91. Retrieved from <http://go.galegroup.com>
- Carless, D., Joughin, G., & Ngar-Fun L. (2006). *How Assessment Supports Learning: Learning-oriented Assessment in Action (Chapter: Improving Assessment, Improving Learning)*. Hong Kong: Hong Kong University Press. Pp. 1-7. Retrieved from <http://www.jstor.org/stable/j.ctt1xwdgd.5>
- Donovan, R, Larson, B, Stechschulte, D., & Taft, M. (2003). *Building Quality Assessment*. 1-49. Action research project, Saint Xavier University and Skylight Field-Based Masters Program. Retrieved from <https://eric.ed.gov/?q=Building+quality+assessment>
- Fox, K., White, S., & Kidd, J. (2011). Program portfolios: Documenting teachers growth in reflection based inquiry. *Teacher and Teaching: Theory and Practice*, 17(1)149-167. Retrieved from <http://www.tandfonline.com/doi/abs/10.1080/13540602.2011.538506>
- Gijbels, D., Watering, G., Dochy, F., & Bossche, P. (2005). Integrating assessment tasks in a problem- based learning environment. *Evaluation in Higher Education*, 30(1): 73-86. Retrieved from <http://www.tandfonline.com/doi/abs/10.1080/0260293042003243913>
- Hago, O. E., & Ali, M. H. A. (June 2015). Assessing English Syntactic Structures Experienced by Sudanese Female Students at Secondary Schools, (2013-2014). *Arab World English Journal (AWEJ)*, 6(2): 219-241. Retrieved from <http://awej.org/images/AllIssues/Volume6/Volume6number2June2015/17.pdf>
- Hamza, Mohammed & Soman, Ahmed (2012). The extent to which the teachers of Jordanian public schools use authentic assessment, their attitudes towards it, and the obstacles to its use. *Journal of Hebron University for Research*, 7(1): 265-283. Retrieved from <http://www.hebron.edu/journal>
- Keyser, S, Howell, L. (2008). *The state of authentic assessment*. Retrieved from [http://www.eric.ed.gov/.content delivery/servlet/ ERICServlet? Accno \(ED503679\)](http://www.eric.ed.gov/.content%20delivery/servlet/ERICServlet?Accno%20(ED503679)).

- Lanting, S. (2000). *An Empirical Study of District-Wide k-2 Performance Assessment Program: Teacher Practices Information. Gained and use of Assessment Resats.* (Published doctoral dissertation, University of Illinois at Urbana champaign, USA. Retrieved from <https://eric.ed.gov/?id=ED452216>
- Marzano, R. (2002) .A comparison of selected methods of scoring classroom assessment. *Applied Measurement in Education, 15*,249-267. Retrieved from <http://www.tandfonline.com/doi/abs/10.1207/S15324818>
- Mueller, J. (July 2005). The Authentic Assessment Toolbox: Enhancing Student Learning through Online Faculty Development. *Journal of Online Learning and Teaching, 1*(1). n.p. Retrieved from http://jolt.merlot.org/documents/vol1_no1_mueller_001.pdf
- Oladele, O. (2011). Knowledge and Utilization of Assessment Techniques by Lecturers in Botswana College of Agriculture. *NACTA Journal, 55*(1): 63-67. Retrieved from <https://www.nactateachers.org/attachments/article>
- Stiggins, R. J. (1987). *Teaching and Learning Community in Unitech.* Retrieved from <http://tlcommunityunitec.ning.com/profiles/blogs/authentic-assessment>. p. 34
- McDonald, J. P. (1992). Dilemmas of planning backwards: Rescuing a good idea: *Teachers College Record, 94*, 152-169. Retrieved from www.academia.com
- Wiggins, G. (1993). *What is Authentic Assessment?* Retrieved from <http://jfmuller.faculty.noctrl.edu/toolbox/whatisit.htm>
- Zu'bi, A. A. (July 2013). Knowledge and Utilization Degree of Authentic Assessment Strategies and Tools by Mathematical Teachers. *Journal of the Islamic University for Education and Psychology, 12*(3): 165-197. Retrieved from <http://www.iugaza.edu.ps/ar/periodical>