

Received on (29-12-2019) Accepted on (21-03-2020)

Science Teachers' Employment of Alternative Assessments for Gauging Students' Learning

Main Researcher

Dr. Ali Khaled Bawaneh

Second Researcher (2):

Prof. Ahmed Boudjema Moumene

University Name & City :

Deanship of Academic Development, Imam
Abdulrahman Bin Faisal University, Saudi Arabia

* Corresponding author:

Deanship of Preparatory Year and
Supporting Studies, Imam Abdulrahman Bin
Faisal University, Saudi Arabia

E-mail address:

akbawaneh@iau.edu.sa

<https://doi.org/10.33976/IUGJEPS.29.1/2021/29>

Science Teachers' Employment of Alternative Assessments for Gauging Students' Learning

Abstract:

The study aimed at gauging science teachers' usability of alternative assessments in light of four demographic variables: teacher specialization, school level, experience, and scientific qualification. The population consisted of all science teachers (physics, chemistry, and biology) in Bani Kenanah Province, Jordan, and the sample comprised 163 teachers. The researchers developed a questionnaire and verified its validity and reliability for collecting data. Some statistical analyses such as Mean, standard deviations and ANOVA tests were conducted to analyze the questionnaire. The results indicated that the degree of using alternative assessments by science teachers for evaluating students' learning is medium in general but with varying degrees from one item to another. The results also revealed that significant differences existed at the level of ($\alpha = 0.05$) for the variables of teacher specialization and experience. However, the results did not show significant statistical differences at the level ($\alpha = 0.05$) of the use of alternative statements for both school level and teacher qualification. The study included several recommendations; the most important of which is the necessity of dispensing prompt instruction on alternative assessments to train teachers adequately and to supplement them with lucid guides to boost their competences and skills in utilizing these innovative models.

Keywords: Teaching & Learning, Alternative Assessment, Science Teachers, and Students' Learning

توظيف معلمي العلوم للتقويم البديل لقياس عملية تعلم الطلبة

المخلص:

هدفت الدراسة إلى قياس مدى قابلية معلمي العلوم لاستخدام التقويم البديل في ضوء المتغيرات الديموغرافية الأربعة: تخصص المعلم، ومستوى المدرسة، والخبرة، والمؤهل العلمي. تألف مجتمع الدراسة من جميع معلمي العلوم (الفيزياء والكيمياء والبيولوجيا) في لواء بني كنانة بالأردن، وكانت العينة 163 معلماً ومعلمة. طور الباحثان استبياناً وتحققا من صدقه وثباته لجمع البيانات. أجريت بعض التحليلات الإحصائية مثل المتوسط الحسابي والانحرافات المعيارية واختبار ANOVA لتحليل البيانات. أشارت النتائج إلى أن درجة استخدام التقويم البديل من قبل معلمي العلوم لتقويم تعلم الطلاب متوسطة بشكل عام ولكن بدرجات متفاوتة من فقرة إلى أخرى. أظهرت النتائج أيضاً وجود فروق ذات دلالة إحصائية عند مستوى ($\alpha = 0.05$) لمتغيرات تخصص المعلم وخبرته. ومع ذلك، فإن النتائج لم تظهر فروق ذات دلالة إحصائية عند مستوى ($\alpha = 0.05$) لاستخدام التقويم البديل لكل من المستوى المدرسي ومؤهل المعلم. وخلصت الدراسة بعدة توصيات؛ أهمها ضرورة توفير برامج تدريبية متخصصة في استراتيجيات التقويم البديل لتدريب المعلمين بشكل كاف وتزويدهم بأدلة واضحة لتعزيز كفاءاتهم ومهاراتهم في استخدام هذه النماذج المبتكرة.

كلمات مفتاحية: التعليم والتعلم، التقويم البديل، معلمو العلوم، وتعلم الطلبة.

Introduction

The alarming low scores obtained by Jordanian students each year draw the attention of stakeholders and instigate action research for approaching the causes of failure which may be due to teaching, learning or testing. How many moans and groans were expressed by parents concerning the school failure of their children without being able to verbalize their effective signs of anguish? In this abysmal, Jordan has noticed local and international efforts to improve students' performance. Yet, the results of the National Assessment of Knowledge Economy Skills Study in 2014 (Ababneh, Al-Tweissi & Abu Lebda, 2015) indicate that more than one-third of the students in grades 5 and 9 are at 'low' and 'very low' levels in the 21st-century skills and knowledge economy skills. This problem further exacerbates among eleventh graders with more than half of the record 'low' and 'very low levels'. The results of the same study revealed that the average performance of students in knowledge, application, and thinking decreases through the progress of students in grades. The average performance of students in the fifth grade (11 years) is better in all aspects of knowledge, application, and thinking. Then, comes the average of the ninth-grade students (15 years). Finally, comes the average performance of students in the eleventh grade (17 years). These findings also apply to the skills of students in solving problems. The results are consistent with other study variables and show that the average performance of students varies according to scientific content and grade. They also indicate differences in the performance of students in the same grade according to scientific content: physics, chemistry, and biology.

In the same context, the results of the International Study of Mathematics and Science in 2015 (TIMSS, 2015) showed that the ranking of Jordanian students was very low. Jordan ranked 32 out of 39 participating countries at a rate of 426, with 60 scores less than the international average: 486. On the other hand, there were differences in the average performance of Jordanian students according to content. The average of chemistry was 438, physics was 424, and biology was 420. As for Jordanian science teachers, TIMSS (2015) reported that 83% of science teachers in Jordan had only a bachelor's degree and that they had about nine years of experience. Students taught by teachers whose experience varied between 10-20 years achieved the highest average performance in science. It is surprising to see in the results of the International Study (TIMSS, 2015) that 62% of science teachers avow that school does not consider student achievement.

The results of the Program for International Student Assessment (PISA) (2015) corroborated those provided by TIMSS (2015). According to PISA (2015), Jordan ranked 63 out of 72 countries. These low results neither meet the expectations nor match the efforts of all the teaching and learning instances in Jordan. This review uncovers unsatisfactory results for the average performance of Jordanian students in science. It indicates a decline at all levels of performance and higher-order thinking. Indeed, the performance of Jordanian students in science does not exceed the knowledge of certain facts and concepts and the deduction of some information from data represented in linear graphs and charts. This level does not correspond to the world's demand for high skills and competitive talents. Hence, this fearful retreat must be controlled to promote students to international competition standards. This conundrum is confirmed by the national report of the PISA study (2015) conducted by the National Center for Human Resources Development (Ababneh, Al-Tweissi, and Abu Lebda, 2017). The report recommended the importance of reviewing evaluation methods used by teachers and the elaboration of a comprehensive and integrated learner-centered approach to improving students' higher-order thinking and problem-solving skills. It also called for the adoption of new alternative assessments to ensure efficiency and reliable results.

Alternative assessment is a recent trend in testing that comes in reaction to the rising problems of traditional assessment in the teaching and learning process (Hamid, 2013). This innovative approach considers assessment as an essential part of the educational system. It lays great emphasis on solving real problems that envisage students in the classroom as well as in their daily life. It also enhances their higher-order thinking skills and subsequently improves the quality of education (Wikstrom, 2007).

The study of Diab (2011) stressed the need to keep pace with innovative learning approaches. This requires that learning should be based on knowledge of the learner and his ability to employ it flexibly in problem

solving and decision-making. Al-Omari and Shehadeh (2010) and Afana (2011) consider that alternative assessment is a new trend and a logical shift to jettison traditional practices based on learners' regurgitation of knowledge. Indeed, the traditional assessment does not reflect students' ability to solve problems, make decisions and possess the 21st-century skills. Other researchers (Napoli & Raymond, 2004; Al-Bashir & Barham, 2012) emphasize the importance of implementing alternative assessments that reflect student learning level and valid quality output of the educational process.

Problem Statement

Teachers spend considerable time and effort in planning and delivering scientific courses to make learning outcomes compatible with teaching strategies. They also select appropriate activities that suit the diversity of students and contents. However, when the fruit-reaping phase comes for students, the teacher quickly designs a test without considering the choice of appropriate assessment methods. Consequently, students get low grades, and teachers start wondering about these unexpected results despite their great-deployed efforts ignoring the models used to evaluate students' level. This ingrained practice is due to teachers' ignorance of novel methods of evaluation (Al-Askah & Al-Doulat, 2016). This is confirmed by the study of Ababneh et al. (2015) published by the Jordanian National Center for Human Resources Development which pointed out that 51% of Jordanian science teachers attributed low grades to students. Shaikh et al. (2007) expressed an urgent need for teachers to learn more about modern alternative assessments to cater to students' needs, thinking skills and learning outcomes.

The most important justification for conducting this study resides in the existence of several inadequacies of the prevalent evaluation tools in science (Al-Askah & Al-Doulat, 2016). Besides, this study is in line with pedagogical development that focuses on the need to review the evaluation systems. The constructivist view of evaluation emphasizes learner knowledge, problem-solving, decision-making, post reflective practice and metacognitive learning (Diab, 2011). Henceforth, the present researchers attempted to investigate the employment of alternative assessments used by Jordanian science teachers.

Study Questions

Main question

- Do Jordanian science teachers employ alternative assessments for gauging students' learning?

The sub-questions are:

1. To what extent do Jordanian science teachers use alternative assessments to evaluate students?
2. Are there statistically significant differences at the level of significance (0.05) on the degree to which Jordanian science teachers use alternative assessments due to the variables of experience, specialization, school level, and qualification?

Objectives of the study

This study emanates from the necessity of utilizing alternative assessments for evaluating students in science streams. Henceforth, the purpose of the present study is to investigate the extent of Jordanian science teachers' employment of alternative assessments for gauging students' learning. Specifically, the objectives of this study sought to investigate:

1. The extent of Jordanian science teachers' use of alternative assessments to evaluate students.
2. The statistically significant differences to which Jordanian science teachers use alternative assessments in terms of four variables: experience, specialization, school level, and qualification.

Importance of the Study

1. Alternative assessments are believed to provide teachers, educators, and decision-makers with reliable information about the level of students' learning and achievement.
2. This study helps develop a road map to improve assessment practices in testing science students.
3. It ultimately offers some recommendations for using alternative assessments to the Ministry of Education, the Directorate of Curricula, and the National Center for Human Resources Development, the Economic and Social Council, and to all the institutions in charge of dispensing teachers' training.

Operational Definitions:

Alternative assessment: is a blanket term used to describe a variety of alternatives to what is considered a standardized form of testing. These innovative trends of evaluation are adopted by the Ministry of Education in Jordan to measure students' performance in real situations. They include performance-based evaluation, observation, communication, peer evaluation, self-evaluation ... etc.

To what extent do Jordanian science teachers use alternative assessment strategies: Teacher questionnaire is designed to measure the degree of instructors' use of alternative assessments.

Science teachers: Teachers who teach various subjects of science (General Science, Physics, Chemistry, Biology, Geology) under the umbrella of the Jordanian Ministry of Education.

Literature Review

Alternative assessment is an umbrella term that covers all types of innovative classroom assessment practices that form an alternative to the traditional approach of standardized assessment or testing. It is also known as authentic assessment, integrative assessment, and holistic assessment. In this context, Al-Mahrooqi and Denman (2018) consider alternative assessment as a 'blanket term' that includes a range of alternatives to 'standardized' exams. Alternative assessment is neither a vagary of fashion nor a fad; it is rather a strong reaction against the recurring failures of the monolithic conventional approach and its persistent hegemony on the field of education since the appearance of the grammar-translation method for teaching and testing. It is a formative assessment that lays great emphasis on the student learning process and continuous development over a specific period according to curricular objectives. Nevo's (1995: 94) provides the following succinct definition of alternative assessment:

In an alternative assessment, students are evaluated based on their active performance in using knowledge in a creative way to solve worthy problems. The problems have to be real problems, i.e., non-routine multi-faceted problems with no obvious solutions. These must be authentic representations of problems encountered in the field of study or the real life of adults.

It goes without saying that there would be no efficiency in education without rethinking the place accorded to teaching, learning, and assessment. Reeves (2000) states that the dissatisfaction with national assessments that aim at measuring the retention of knowledge by pen and paper led to the emergence of alternative assessments to evaluate higher-order attainments that involve profound comprehension and employment of knowledge in complex, realistic contexts. In this perspective, Oliver (2019: 1) strongly articulates that the triangle of effective education comprises three equally important pillars: student-centered teaching, blended learning, and transformative assessment. According to Gozuyesil and Tanriseven (2017), recent trends encourage the use of alternative assessment tools in class in line with the recommendations made by the updated curricula. Contrary to traditional evaluation, which examines students on a cumulative set of work for a specific period, alternative assessment adopts a holistic approach for testing that allows students to prove their real learning capacities. Traditional assessment proves to be disadvantageous for learners for its compression of a long period of learning into a short time of testing. This common practice does not reflect students' level at all, but it really represents a blatant injustice to their legitimate rights for success.

Several types of alternative assessment came to the scene (Brown and Hudson, 1998; Zimbicki, 2007; Rodomanchenko, 2017) including portfolios, diaries, rubrics, journals, Knowledge mapping, oral presentation, demonstration, interaction, simulation, role-play, discussion, performance, confrontation, worksheets, short tests, unit-end tests, open-book tests, automatic/systematic observations, summaries, communication, crib Sheets, conferences, interviews, collaborative tests, group-work, self-evaluation, peer

evaluation, open tasks, projects, monitoring lists (write-off), estimation scales, narrative records, and take-home exams. Indeed, the list is no longer exhaustive.

Alternative assessments seem to yield many benefits in the field of education (Borko et al. 1997; Gozuyesil & Tanriseven, 2017; Al-Mahrooqi & Denman, 2018). Firstly, they strive to consider the students' needs, styles, creativity, and involvement. Secondly, they cater not only to students' declarative knowledge but also for their procedural knowledge as well. Contrary to traditional evaluation that is limited in time and content through the provision of a few questions that examinees must answer during a short period of time in a final test, alternative assessments do not fail to cover all the breadths and depths of the curriculum design. Besides, the reduction of learning efforts in a short product exam does not reflect students' true capacity, but it generally results in the reception of low grades. In addition, alternative assessments endeavor to examine applied proficiency rather than theoretical knowledge. They also emphasize promoting reflective thinking and problem solving rather than regurgitating facts and knowledge. This situation leads many students to display an utter lack of interest in studying and persevering and therefore escalates the rate of their dropouts. The effective implementation of workshops on innovative evaluation models would enable teachers to familiarize themselves with them and consequently use them in real classroom situations.

non-experienced teachers may show resistance to change for adopting alternative assessments. Their reluctance to change may be due to mere ignorance about the issue. Also, teachers seem to be very reluctant to get rid of old practices because some unfounded assumptions played a role in making them prefer traditional assessment. For example, Jordanian teachers usually think that alternative assessments are too demanding to grasp and difficult to prepare; thus, they stick to traditional assessments believing in its suitability for large classes and saving time. They also ruminate that they are time-consuming for preparation and more difficult to grade than conventional exams. Certainly, this assertion may hold true to some extent for some formats but not for all. Above all else, if the results can be rewarding, the endeavor will be well worth the effort to achieve effective education. In this context, Taf, Dorf, and Lazarowitz (2000) used a project-based alternative assessment system and attested that this assessment is 'complex', but "it proved to be meaningful and suitable" (p. 187). They also reported that "The project-based curriculum and its associated assessment system require an investment of time and effort not only of teachers but of parents and community members as well" (p. 187).

Moreover, some stakeholders seem to be suspicious about alternative assessments admitting that they are not yet well-experimented upon. In this direction, Brown, and Hudson (1998) notify the pejorative connotations that wrap up alternative assessments and hamper their acceptability and enactment including the claim that they are untried and unsupported by theoretical and empirical research. Almqati (2015), on his part, renders the constraints that prevent the use of alternative assessments to the curriculum, the teacher himself, the learner and the educational environment. Eventually, there is still a pressing demand to carry out further experimental studies on alternative assessment models.

Undoubtedly, holding regular professional development training can help teachers become familiar with alternative assessments. In addition, reflective teaching practices should be encouraged amongst teachers to enable them to uncover their unfounded assumptions and to keep abreast of recent professional developments. In line with this, Borg (2003: 81) strongly admits, "Teachers are active, thinking decision-makers who make instructional choices by drawing on complex, practically-oriented, personalized, and context-sensitive networks of knowledge, thoughts, and beliefs". Teachers are likely to change and adopt novel teaching methods and assessments if they are engaged in a constructive dialogue with academics and researchers. In this context, Borko, Mayfield, Marion, Flexer, and Cumbo (1997) designed a project to help third grade teachers develop and implement 'classroom-based performance assessments' in mathematics and literacy. The results reveal that conversations between teachers and researchers prove to be an efficient way for assisting teachers to get rid of their ingrained past practices and adopt innovative methods.

Most importantly, the last two decades noticed an unprecedented concern with theoretical and empirical research testing out the real impact of alternative assessments on students' learning achievement and

teachers' attitudes. A battery of studies was carried out to explore one alternative assessment format and its impact on one or several dependent variables. For instance, Sandford and Hsu (2013) found that alternative assessments and especially portfolios deserve reconsideration because of their invaluable and motivating role in both formative and summative evaluation. In the same perspective, Gozuyesil and Tanriseven (2017) also demonstrated that portfolios as an alternative assessment technique have a substantial impact on students' academic achievement revealed especially in mathematics courses. Jangi and Farjami (2017) investigated the effect of 'collaborative assessment' on students' classroom participation and development of a 'sense of classroom community' (SCC) in face-to-face classes. Considering the positive results, they endorsed a collaborative assessment to enhance SCC. In another direction, Rodomanchenko (2017) investigated 'roundtable discussions' for EFL/ESL proficiency levels. She enumerated their advantages in furthering face-to-face interviews and honing academic skills and consequently recommended them for both summative and formative assessment. Alternatively, Chicaiza (2018) concluded that the use of the problem-based approach in the Moodle platform improved considerably students' oral production. The results of these studies show the positive effects of alternative assessments and unveil the myths surrounding their practical contribution.

Many experimental studies were conducted to investigate an ensemble of alternative assessments and their effects on several dependent variables. In her doctoral action research study, Zimbicki (2007) examined the impact of a range of alternative assessments on 7th-grade students' motivation and self-efficacy. The testing comprises six alternative assessments: performance, cooperation, product, interaction, portfolios, and oral testing. The results showed that students reached higher levels of motivation, self-efficacy, confidence, sense of accomplishment and freedom. Alnajim (2013) also proved the efficiency of alternative assessments for honing self-efficacy and autonomy. In addition, Fook and Sidhu (2014) investigated assessment practices in an institute of higher learning in the US. They found that students preferred formative assessment to the final examination and expressed positive attitudes toward constructive feedback from lecturers for better learning. Likewise, Kirikkaya, and Vurkaya (2011) found that using alternative assessments improved not only students' achievement but also their attitudes towards learning and reducing anxiety. For the same reason, Juhaina and Nasser (2018) explored the training of prospective and practicing mathematics teachers in alternative assessment and its impact on their attitudes toward alternative assessments in mathematics. While the two groups differed in their attitudes toward alternative assessment before the course, they expressed constructivist beliefs and positive standpoints toward these tools after taking part in the course.

Concurrently, and more importantly, a plethora of studies revealed the importance of alternative assessments for developing teaching and learning. Hamid (2013) showed that alternative assessments played a great role in developing students' reflective thinking. Similarly, Almazroue (2009) and Bramwell and Rainford (2014) concluded that alternative assessments developed cognitive concepts and skills. Other researchers (Al-Bashir and Barham, 2012; Alsuwaihiri, 2013; Hassan, 2013) found that alternative assessments not only helped students to gain scientific skills and concepts but also to connect what they learned in the classroom to their daily life as well. All these studies proved the effectiveness of alternative assessments for improving valid and reliable evaluation.

In a similar line, Abu Libdeh et al. (2011) aimed at uncovering the skills that are mastered and not mastered by the students of the basic stage in Jordan in the light of the results of the international study of student evaluation Pisa (PISA, 2009). The results indicated that the students were more successful in answering first multiple-choice questions, followed by the answers to the closed options that required specific answers available directly in the text. The students perform less in the items that require higher mental processes such as analysis, synthesis, and evaluation than in the answer rates of other types of items. Jordanian students' general performance at lower levels of thinking was better than that of their upper levels of thinking. The results also showed that female students performed better than males on all forms of items (multiple-choice questions, closed and open questions). The recommendations of this study point out the need to train teachers on how to construct test items that include life situations and measures as well as

higher-level skills needed by students in their daily life. The net result of these studies is that Jordanian teachers need urgent knowledge and practice about alternative evaluation strategies.

Several studies attempted to uncover the degree of employment of alternative assessments by science teachers. Alasqa and Aldulat (2016) examined science teachers in Saudi Arabia to find out the degree of their employment of alternative evaluation and its impact on female students' performance. The findings indicated that the degree of employment is generally moderate irrespective of teachers' experience, job, and qualification. Henceforth, the study recommended the need to change the ethos of teachers' vis-a-vis alternative assessment and exhorted the authorities to organize workshops on alternative assessments for the teaching staff and to implement this kind of evaluation in all scientific curricula. In line with this, and in order to identify the level of teacher satisfaction of using alternative assessments, Al-Hjaili (2016) targeted 2617 Saudi teachers from most of the regions of Saudi Arabia. The results indicated that the average satisfaction of teachers with the use of alternative assessments ranged from 1.95 to 2.49. The net recommendation stressed the need to change the attitude of teachers, students, and parents about alternative assessment and to provide various incentives for teachers who use these methods for examining their students.

In the same direction, Al-Humaidi and Al-Safairi (2016) aimed at determining the extent to which Arabic language teachers use alternative assessment techniques. The results indicated that the degree of using alternative evaluation methods was medium with arithmetic mean: 2.89 while the arithmetic average of the assessment-based on pen and paper was high: 4.73. However, assessment-based on observation, performance, and interviews recorded 2.68 - 3.03 averages, but self-based assessment registered just 1.64. The results showed that there were statistically significant differences due to sex in favor of females in the assessments based on performance, observations, interviews, and self-assessment. The results were similar in terms of the statistical significance of the variable of teacher scientific qualification. As regards the variable of experience, the results signposted that there were statistically significant differences due to assessments based on performance, feedback, and interview. Yet, the study did not reach significant statistical differences due to the variable of experience for both pen-and-paper assessment and self-based assessment.

A multitude of studies prospected the repercussions of alternative assessment for ensuring effective education. Odeh (2015) conducted a study to check its impact on the achievement of ninth-grade students of science through an experimental study containing an achievement test, rubric scale, and a rating scale. The results revealed that there were statistically significant differences in the achievement of science for the benefit of the experimental group while there were no differences in the achievement of students according to their gender. The researcher recommended the use of alternative assessments for evaluating students in various subjects because of its positive impact on enhancing students' achievement. Concomitantly, Allouh (2015) studied the impact of alternative assessment on the development of mathematics capacity through the elaboration of tests and note cards to measure female students' mathematical abilities. The results pinpointed significant differences in both tests regarding students' mathematics ability and performance skills for the benefit of the experimental group among the tenth-grade students. Finally, the study recommended the importance of using alternative assessments in learning. Concurrently, Ashkar (2015) investigated the impact of alternative assessment on developing mathematical thinking. He experimented three types of assessment on three different groups: (1) peer assessment, (2) performance evaluation with conceptual maps and (3) conventional assessment. The study found statistically significant differences in mathematical thinking in favor of the second group students who used conceptual maps in students' assessment. In the same perspective, Awawdeh & Nasser (2018) examined the training of mathematics teachers in alternative assessments and their attitudes toward these new methods. The findings signposted substantial changes in their attitudes toward alternative assessment accompanied by a shift toward using ample assessment models. In sum, all these results lend substantial support to the use of alternative assessment practices.

Finally, yet importantly, the general results of the studies mentioned in the literature review reveal the numerous benefits of alternative assessments for students, teachers, and academics by large. Nonetheless, researchers need to conduct more theoretical and empirical research on this delicate issue to demystify the issue of alternative assessments so as it will no longer confuse new practitioners in the field of education.

Method

Population

The population of this study comprised all male and female science teachers enrolled in Bani Kenanah educational province from Irbid Governorate in the 2018/2019 academic year. It was representative and homogenous covering almost all the existing social government teachers in Jordan in terms of gender, nationality, education, culture native language, and exposure to English as a foreign language.

Sample

A sample of 163 Jordanian teachers from Bani Kenanah education province participated in this study. They were selected randomly without any discriminating factor (Gay & Airasian, 2003). All of them held a major in physics, chemistry or biology. While some of them held a Diploma in education or ICT, others held a Master or Ph. D in their specialization in science education. Some of them had teaching experience in the Jordanian Ministry of Education and others in the Gulf countries. All of them had the opportunity to teach heterogeneous classes. Table 1 shows the participants' distribution according to their specialization, school level, experience, and qualification.

Table 1: Participants' Distribution by specialization, level, experience, and qualification.

Table 1 participants' by school level, and The sample science	Variable	Value Label	N	%	shows distribution specialization, experience, qualification. included 163 teachers.
	Specialization	Physics	41	25%	
		Chemistry	72	44%	
		Biology	50	31%	
	School-level	Basic	120	74%	
		Secondary	43	26%	
	Experience	(1-7) Years	54	33%	
		(8-15) Years	47	29%	
		More than 15 Years	62	38%	
	Qualification	BA	91	56%	
		MA	49	30%	
		Ph. D	23	14%	

First, teachers were divided into three groups according to their specialization. The lowest group includes teachers of physics with 25%, the second group forms 31% of teachers of biology, and the third group consists of 72 teachers of chemistry forming the highest percentage: 44%. According to the school level, the sample was divided into two groups. The smallest group represents secondary schools with 26%, and the highest percentage covers 120 basic schools with 74%. Regarding experience, the sample was distributed as follows: 38% of the teachers with more than 15 years' experience, 33% with an experience ranging from 1-7 years, and 29% of the teachers had 8-15 years' experience. As for teachers' qualification level, level one included 91 Bachelor's holders out of 163 with 56% representing the highest percentage, level two included MA holders with 30% and level three comprised Ph. D. holders with 14% only representing the lowest level.

Study Variables

- Independent variables:

1. Teachers' specialization (Physics, Chemistry, and Biology)
 2. School-level (Basic, Secondary)
 3. Teachers' experience per year: (1–7, 8–15, 11–15, more than 15)
 4. Teachers' education/qualification: (Bachelor, Bachelor with a high diploma, MA, and Ph. D.)
- Dependent variables:
 1. The degree of employing alternative assessment in the classrooms by science teachers.

Research Instrumentation

The present researchers reviewed theoretical and empirical research literature related to alternative assessments and their employment for the sake of promoting students' evaluation and providing valid and reliable measurements of their real pedagogical achievements. The researchers adopted the research instrument from Al-Hamidi and Al-Dhafiri (2016) because it covers the largest number of alternative assessment strategies and tools and it is a recent study with less than five years' span. The researchers elaborated it according to the Likert scale with five levels (Very often: 5, Often: 4, Neutral: 3, Rarely: 2, and Very rarely: 1). The tool aimed at measuring the extent of using alternative assessments in the classrooms by science teachers.

Validation of the Instrument

In addition to the present researchers' validation of the test, seven experts validated the initial version of the study instrument. Two of them held a Ph. D degree in curriculum and instruction, and they taught at the university level. Three held an MA degree in science education and worked as educational supervisors at Bani Kenanah education province. Two-science teachers: one of them holds an MA degree in chemistry and the other a Bachelor in physics and a high diploma in teaching science. The tool included 35 items, yet five items (discussion assessment, peer observation, retake policies, replacing test with summaries, and assessment through focus group) were deleted based on the experts' recommendations and the wordings of several items were amended. Overall, the final version included 30 items.

Reliability of the Instrument

To calculate the reliability of the instrument, the researchers applied the tool one time only for 37 science teachers from the second Irbid education province. Then, they calculated the Cronbach Alpha by using the SPSS program, and the reliability coefficient was 0.79. This value was considered acceptable in the social sciences (Gay, Mills, & Airasian, 2009; Al-Kellani & Al-Shraifeen, 2011).

Statistical Standard

The following equation was adopted for classifying items (Al-Rashidi, 2018).

= (Upper limit of scale - minimum scale) / number of required categories

= (5 - 1) / 3 = 1.33

The categories are (1 - 2.33: Weak (W), 2.34 – 3.60: Medium (M), and 3.60 – 5: Strong (S)).

Treatment Procedures

For conducting this study, we asked for a recommendation letter from the Jordanian Ministry of Education. As soon as we received the approval letter, we embarked on piloting the research work.

The next step resided in selecting the sample from the population and designating the participants to collect the data. Then, we proceeded with the random selection of the teachers in coordination with the Provincial Directorate of Education in Bani Kenanah. In early November 2018, we started collecting the data from the teachers, and this process lasted until the end of December 2018. Eventually, the SPSS program was used to analyze the quantitative part of the data, and valid statistical analyses were provided.

Findings

To answer the first question of the study: "To what extent do Jordanian science teachers use alternative assessments to evaluate students?", the present researchers calculated the mean and standard deviation of the instrument items prepared for this purpose, and the results were cast on Table 2:

**Table 2: Means and standard deviation of Alternative assessment among Jordanian science teachers
(N=256; Weak: W; Medium: M; and Strong: S)**

Type of alternative assessments	Mean	SD	Level
1. Performance-based assessment	2.46	1.19	M
2. Presentation-based assessment	3.04	1.26	M
3. Assessment-based on demonstration	3.14	1.46	M
4. Assessment-based on interaction	2.87	1.32	M
5. Assessment-based on simulation and playing role	3.30	1.42	M
6. Assessment-based on discussion and confrontation	3.11	1.45	M
7. Assessment-based on pen and paper	2.18	1.10	W
8. Assessment-based on worksheets	2.69	1.44	M
9. Assessment-based on short tests	2.82	1.28	M
10. Assessment-based on unit-end tests	2.77	1.39	M
11. Assessment-based on monthly tests	2.33	1.21	W
12. Assessment-based on automatic observation	2.91	1.36	M
13. Assessment-based on incidental systematic observation	3.14	1.40	M
14. Communicative-based assessment	2.82	1.29	M
15. Conference-based assessment	4.59	1.37	S
16. Interview-based assessment	4.21	1.40	S
17. Group-work assessment	3.13	1.39	M
18. Assessment-based on questions and answers	2.23	1.18	W
19. Self-revision-based assessment	3.50	1.45	M
20. Self-assessment-based evaluation	3.48	1.35	M
21. Assessment-based on peer evaluation	3.99	1.35	S
22. Assessment-based on student's diary	3.75	1.44	S
23. Assessment-based on student's file	3.49	1.41	M
24. Assessment-based on open tasks	3.90	1.37	S
25. Project-based assessment	4.09	1.33	S
26. Assessment-based on monitoring lists (write-on)	3.07	1.36	M
27. Assessment-based on estimation scales	2.62	1.30	M
28. Assessment-based on oral-estimation scales	2.79	1.40	M
29. Assessment-based on describing the learning process	3.71	1.41	S
30. Assessment-based on narrative record	3.94	1.50	S
Overall	3.20		M

The results in Table 2 show that the overall mean for science teachers is 3.20. This demonstrates that their alternative assessment level in the classrooms is medium and that the highest means for science teachers is 4.59 (strong item). This indicates that science teachers are very active in attending educational conferences and publishing educational research works in compliance with the requirements approved by the ministry of education for ranking them. This was followed directly by item 16 with a calculation mean of 4.21 (strong item) which supported that science teacher was seriously using alternative assessment methods in teaching and learning. Then, comes item 25 with a mean 4.09 (Strong item) in the third rank. This indicates that science teachers used project-based learning and assessment for improving students' competencies. However, the seventh item comes in the last order in terms of the mean 2.18 (weak item). This item reports that science teachers are not trying to employ the traditional assessment based on pen and paper in their teaching and that they are motivated to practice new alternative assessments in their classes. The next

lowest two items 18 and 11 support this with a mean 2.23 and 2.33 respectively (weak items) and demonstrate that teachers do not focus only on using question-answer assessments or monthly or midterm exams to assess students.

To answer the second question: "Are there statistically significant differences at the level of significance (0.05) on the degree to which Jordanian science teachers use alternative assessments due to the variables of experience, specialization, school level, and qualification?", the researchers calculated means and standard deviations as shown in Table (3):

Table 3: Means and standard deviations of teachers' alternative assessment practices in the classrooms according to experience, specialization, school level, and qualification

Variable		N	Mean	SD
Specialization	Physics	41	3.27	0.75
	Chemistry	72	3.04	0.75
	Biology	50	3.39	0.72
	Total	163	3.20	0.76
School Level	Basic	120	3.18	0.80
	Secondary	43	3.28	0.61
	Total	163	3.20	0.76
Experience	(1-7) Years	54	2.86	0.71
	(8-15) Years	47	3.30	0.72
	More than 15 Years	62	3.42	0.72
	Total	163	3.20	0.75
qualification	BA	91	3.09	0.71
	MA	49	3.34	0.85
	PhD	23	3.35	0.66
	Total	163	3.20	0.76

The results in a table (3) showed that there is a difference in the mean of specializations in favor of biology with the highest mean: ($M = 3.39$, $SD = 0.72$) and chemistry with the lowest mean: ($M = 3.04$, $SD = 0.75$). At the school level, the data showed that there is a small difference: (0.10) in favor of secondary schools: ($M = 3.28$, $SD = 0.61$). Regarding teachers' experience, table 3 indicates that the category of teachers with a teaching experience of more than 15 years is the highest one in using alternative assessments in the classrooms with a mean 3.42 and a standard deviation of 0.72. Yet, alternative assessment practices within the category teaching experience: 1-7 years are the lowest with a mean 2.86 and a standard deviation of 0.71. Finally, the results indicate that Ph. D holders' mean is the highest with a mean 3.35 and a standard deviation of 0.66, whereas the bachelor's holders registered the lowest mean 3.09 with a standard deviation of 0.71.

When we examine the above results, we find that there are apparent differences in the mean of the degree of employing alternative assessments in the classrooms by science teachers according to the variables of specialization, school level, experience, and education. To ascertain the validity of the differences, the researchers performed the ANOVA analysis, and the results were presented in Table 4 as follows:

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	Hypothesis	1613.379	1	1613.379	867.739	.001
	Error	3.774	2.030	1.859 ^a		
Specialization	Hypothesis	3.784	2	1.892	3.409	.035

	Error	88.803	160	.555 ^b		
Intercept	Hypothesis	1319.875	1	1319.875	3958.981	.010
	Error	.333	1	.333 ^a		
School level	Hypothesis	.333	1	.333	.582	.447
	Error	92.254	161	.573 ^b		
Intercept	Hypothesis	1645.462	1	1645.462	338.351	.003
	Error	9.740	2.003	4.863 ^a		
Experience	Hypothesis	9.782	2	4.891	9.451	.000
	Error	82.805	160	.518 ^b		
Intercept	Hypothesis	1278.904	1	1278.904	1100.674	.000
	Error	2.701	2.324	1.162 ^a		
qualification	Hypothesis	2.535	2	1.267	2.252	.109
	Error	90.052	160	.563 ^b		

Table 4: ANOVA test of science teachers' alternative assessment practices in the classrooms

Table 4 shows that there are statistically significant differences among science teachers' alternative assessment practices in the classrooms due to two independent variables: specialization ($F= 3.409$, $\alpha < 0.05$) and experience ($F= 9.451$, $\alpha < 0.05$). However, the results indicate that there are no statistically significant differences among science teachers' alternative assessment practices due to the two independent variables: school level ($F = 0.582$, $\alpha = 0.447$) and teachers' qualification ($F = 2.252$, $\alpha = 0.109$).

Discussion and Results

Results of the First Question

The results of the study indicated that the average level of alternative assessments employed by science teachers was average. This reveals that this degree of use neither fulfills the required level expected by the Ministry of Education nor matches the great efforts made for the development of education in Jordan. Some researchers (Almorshid, 2014; Al-Asqah and Dulat, 2016; Al-Sa'ydeh, 2016) attributed the reasons for these results to:

1. The fresh recent use of alternative assessment methods, insufficient knowledge of many teachers, and parents' lack of confidence in the usefulness of this form of evaluation. Parents argue that alternative assessments are neither objective nor consistent in themselves and that the likelihood of teacher bias toward some students seems to be quite omnipresent.
2. Educators believe that alternative assessment needs a longer time for preparation and implementation at a time when teachers suffer from the content load and insufficient time allocated for its coverage. They also attribute this problem to the high economic cost of employing alternative assessment compared to conventional assessment.
3. Al-Hjaili (2016) believes that the heavy tasks allocated to teachers such as academic and administrative works, the limited time of lectures, and the classroom crowdedness reduce the ability of teachers to manage the classroom.
4. Other factors also aggravated the situation such as the lack of financial incentives for outstanding teachers and the use of conventional methods instead of alternative assessments. Several researchers (Ostaz, 2011; Hassan, 2013; BuQhoos, 2017) point to several justifications that may explain this finding.
5. Additionally, the scientific content of science courses is not designed in symbiosis with alternative assessment and therefore does not encourage science teachers to adopt this strategy.

Results of the Second Question

The second question tackles the level of usability of alternative assessments by science teachers in terms of teaching experience, specialization, educational level, and school level. The reasons that may have led to the existence of differences of statistical significance for the variables of teacher experience and specialization were discussed. Then, the reasons related to the absence of significant statistical differences

related to the variables of teacher educational qualification and school level (basis or secondary) in teaching sciences were highlighted.

With regard to teaching experience and specialization, the results showed that the highest use of alternative assessments was achieved by teachers with a long experience (15 years and beyond), then intermediate experience (8-15 years), and finally novice teachers (1-7) years. This result sounds reasonable for this variable (Al-Sharaa and Zaza, 2013; Al-Humaidi and Al-Safairi (2016); Ryan, 2014):

1. The most experienced teachers often use alternative assessments because they attended various professional development workshops during their careers and developed knowledge of using various types of suitable alternative assessments for testing.
2. In addition, more experienced teachers are more knowledgeable about scientific content and thus are more familiar with choosing the best ways to evaluate targeted learning outcomes.

The results showed that there were statistically significant differences due to teacher specialization. The highest use of alternative assessment was in favor of Biology teachers, then Physics teachers and finally Chemistry teachers. This may be due to:

1. The nature of the biology content that contains many activities and practical experiments. Similarly, physics requires the use of alternative assessments to evaluate students' performance rather than relying solely on pen and paper tests. As for the content of chemistry, it also contains experiments and activities that require the activation of alternative assessments. On the one hand, this may not be possible because many activities and experiments have not been carried out due to the lack of tools and equipment necessary for this.
2. On the other hand, the laboratories are not properly equipped and do not comply with the rules of public safety.
3. In addition, many schools, especially those which are far from the centers of cities, are found to lack laboratories at all. Therefore, teachers only provide the scientific content in theory and make recourse to traditional assessment.
4. It is important to note that in many basic schools, the teacher of chemistry can be a physiologist or biologist. He is charged with this task in order to complete his teaching load because he is considered as a specialist in science and that is closer to chemistry. Of course, this reduces his potentiality of using alternative assessments because he is not a specialist, especially in practical experiments which greatly need to galvanize alternative assessments.

Concerning the variables of teacher educational credential and school level, the results indicated that there are no statistically significant differences in the level of science teachers' use of alternative assessment strategies and tools due to their educational attainment and school level (primary or secondary). Many researchers (Abu Sultan and Abu Asker, 2017; BuQhoos, 2017) believe that:

1. The meager use of alternative assessment strategies by science teachers can be rendered to other motives that are not related to scientific qualification or school level.
2. Alternative assessment is a new trend, and many educational supervisors still focus on using the traditional pen and paper assessment to measure students' learning progress (Al-Humaidi, Al-Shafairy, 2016; Al-Rifaa'i, Tawalbeh, and Al-Qaood, 2012). Therefore, regardless of their scientific qualifications, teachers did not receive any training on alternative assessments.
3. There are no significant differences in the material income according to the difference of experience and scientific qualification as well as the difference in educational levels. This, of course, reduces the inspiration of science teachers to devote them to experimenting on modern teaching methods and evaluation strategies.
4. On the other hand, science teachers follow the same professional development programs that are offered free of charge by the Ministry of Education both before and during the service irrespective of their academic qualification or school category.

Eventually, Jordanian teachers, in general, suffer from the lack of time and the heavy academic and administrative burdens added to the difficulty of economic and social conditions. With the rising cost of

life, teachers feel the need to challenge hardship and to find new sources of income after the end of school hours to meet their family provisions. As a result, teachers return home very tired at the end of the day (Al-Rfou'a, 2017 Al-Sa'ydeh, 2016; Ostaz, 2011). This situation reduces the chances of good preparation and diversification of teaching methods and assessment mechanisms.

Conclusions

The article examines the rationale behind the national shift from traditional to alternative assessment practices. It is underlined that students' low grades in sciences led teachers to doubt about the efficiency of the traditional evaluation tests and to look for new assessment tools as an alternative. This was boldly expressed by TIMSS (2015), PISA (2015), and the National Assessment of Knowledge Economy Skills Study in 2014 (Ababneh, Al-Tweissi & Abu Lebda, 2015).

Henceforth, to diagnose the causes of the alarming falling standards, the article displays the philosophical foundations and theoretical assumptions of traditional and alternative assessment and their consecutive effects on students' achievement. Traditional assessment involves the average grading of a cumulative set of work in a mid-term or final exam. Teachers usually make recourse to this approach for its expedience to save timing in grading large size classes. However, alternative assessment is considered as a holistic method of grading student performance enabling them to demonstrate all the scope of what they have learned rather than being constrained to a set of questions to be answered in a limited period. The success of using alternative assessments depends not only on teachers' resources and understanding of this broad approach but also on students' comprehension and familiarity with its variegated models as well.

It is worth mentioning that most if not all current studies (Gozuyesil and Tanriseven, 2017; Chicaiza (2018) Oliver, 2019) including the present one enumerated the benefits of alternative assessments and accentuated their relevance to the evaluation of students' progress and performance through their learning process. Innovative assessments not only serve to grade students but also to assist in learning factual knowledge, professional skills, critical reflection, and problem-solving as well.

The main aim of the study was to explore the level of employing alternative assessment tools by science teachers in teaching Jordanian students in basic and secondary schools, namely in Bani Kenana province, Jordan. The results showed that the employment rate of using alternative evaluation by teachers is generally average. However, the findings revealed the existence of different degrees of employment between the various scientific disciplines experimented upon. The results also indicated that there are statistically significant differences in the degree of use of alternative assessment strategies by science teachers according to the two study variables: teacher specialization and teacher experience. Nonetheless, no significant statistical differences are signaled concerning teacher scientific qualification and school level (basic or secondary) where they work.

The findings of this article corroborate several studies that sought to examine the degree of utilizing alternative assessments by science teachers. The results uphold Alasqa and Aldulat's (2016) study that found that the degree of employment of alternative assessment by science teachers in Saudi Arabia on female students' performance is generally 'moderate' regardless of teachers' experience, job and qualification. Our findings also support Al-Humaidi and Al-Safairi (2016) study that aimed at detecting the level of employment of alternative assessment by Arabic language teachers and showed that their use of these innovative models was 'medium'. Additionally, the results of this article sustain Al-Hjaili (2016) work that targeted 2617 teachers from most of the regions of Saudi Arabia and found that the average satisfaction of teachers with the use of alternative assessment ranged from 1.95 to 2.49.

Recommendations

This study underlines the need for training teachers in using alternative assessments for ensuring effective learning and fair student evaluation. This tendency goes in line with the bulk of recent research literature. For example, Odeh (2015) recommended the use of alternative assessments for evaluating students in various subjects because of its positive impact on enhancing students' achievement. Concurrently, Allouh (2015) endorsed the importance of using alternative assessments in learning. Alongside, Ashkar (2015) found significant improvement in student mathematical thinking. So, due to the positive results of this kind

of assessment on students and the moderate use of them by teachers, adequate teacher training should be implemented to familiarize both teachers and students. If teachers are well trained in this field, their professional development would be developed, and their contribution would be enhanced. This was also proved by Awawdeh & Nasser (2018) who examined the training of mathematics teachers in alternative assessments and their attitudes toward these new methods. Results signposted substantial changes in their attitudes toward alternative assessment accompanied by a shift toward using ample assessment models. Undeniably, alternative assessments should not be applied haphazardly without appropriate training and deep reconsiderations; they rather require effective reflection for ensuring sound implementation.

Considering the theoretical and empirical positive findings, it is believed that it is important to encourage teachers to activate their use of alternative assessment strategies. Officials should aid teachers to keep abreast of recent evaluation developments and provide them with all the financial and moral support they need. There is an urgent necessity to design and implement specialized training programs in the mechanisms of using classroom alternative assessments for all teachers at all levels. It is also important to familiarize students with various alternative assessment models to be ready for any evaluation in order to ensure long-life learning.

In sum, this study consists of some limitations. One of these resides in the human factor. The study covered only science teachers in the education directorate of Bani Kenana, Jordan. Further research needs to be carried out at a large scale in Jordan to determine the efficiency of alternative assessments. Another limitation may be inherent in the research tool used in the study and its impact on the degree of reliability and validity of the questionnaire. Additionally, this study did not utilize any other research instrument to scrutinize the real attitudes of both students and teachers in this experiment. Getting ideas about perceptions seems to be of paramount importance in this type of research. It would also be interesting to further research into one alternative assessment model compared with traditional testing to demonstrate which evaluation accurately reflects students' real learning and to highlight their positive and negative outcomes.

References

- Ababneh, E. Al-Twaissi, A., & Abu Lebda, K. (2015). Program for International Student Assessment (PISA2015): *National Report. National Center for Human Resources Development (NCHRD)*, (1787/4/ 2015), Amman, Jordan.
- Ababneh, E., Al-Twaissi, A. & Abu Lebda, K. (2017). Program for International Student Assessment (PISA2015): *National Report. National Center for Human Resources Development (NCHRD)*, (6350 /12 / 2017), Amman, Jordan.
- Abu Lebda, K., Al-Twaissi, A. & Ababneh, E. (2017). Jordanian National Report of TIMSS 2015 Study. *National Report. National Center for Human Resources Development (NCHRD)*, (2017/8/3881), Amman, Jordan.
- Abu Lebda, K., Hijjazeen, N., Hamed, Sh., and Qudah, Kh. (2011). Detection of the Mastered / Not Mastered Skills by the Students in Upper Basic Cycle in Jordan in light of the results of the Program for International Student Assessment (PISA 2009). *National Report. National Center for Human Resources Development (NCHRD)*, (2017/8/3881), Amman, Jordan.
- Abu-Sultan, A; and Abu Asker, M. (2017). The Reflective Practices of Science Teachers in the Upper Basic Stage in northern Gaza. *Journal of Palestine University for Research and Studies*. 7(1), 1-27.
- Afana, M. (2011). The Reality of Using Arabic Language Teachers for Assessment Techniques in the Preparatory Stage in UNRWA Schools in the Gaza Strip in the Light of Recent Trends (Unpublished Master Thesis), Faculty of Education, Islamic University, Gaza.
- Afana, A., & Abu Mallouh, M. (2005). The effect of a proposed model for the treatment of misconceptions of the mathematics concepts among low achievement students in the seventh grade primary Gaza. *Paper presented at the 2nd Educational Conference, "Palestinian children between the challenges of reality and the aspirations of the future"*, Educational Faculty. November 22-23, Islamic University, Gaza, Palestine.

- Al-Askah, H., & Al-Doulat, A. (2016). The Degree of Using Alternative Assessment Methods by Intermediate Stage' Female Science Teachers in Al-Qasem District in Saudi Arabia. *Studies in Educational Sciences*. 43(1), 37-48.
- Al-Bashir, A., & Barham, A. (2012). Using alternative assessment strategies in assessing students' learning in mathematics and Arabic in Jordan. *Journal of Educational and Psychological Sciences*. 13(1), 241-270.
- Al-Hamidi, H. & Al-Dhafiri, M. (2016). The extent of using Arabic language teachers in the middle school in Kuwait of alternative assessment methods from their point of view. *Educational Sciences*. 3(1), 167-210.
- Al-Hjaili, M. (2016). Degree of satisfaction and use of teachers in Saudi Arabia for alternative assessment methods in evaluating the teaching process in general education and the challenges to its use. *Educational Sciences*. 2(1), 205 -261.
- Al-Kellani, A. & Al-Shraifeen, N. (2011). *Introduction to research in educational and social sciences* (3rd Ed.). Amman, Jordan: Dar Almaisarah for publishing, distribution and printing.
- Allouh, Sh. (2015). *The Impact of Employing the Alternative Evaluation on Developing the Mathematical Ability among Female Tenth Graders in Gaza*. (Unpublished Master Thesis), Faculty of Education, Al- Azhar University, Gaza.
- Al-Mahrooqi, R. I. and Denman, C. J. (2018). Alternative Assessment. *The TESOL Encyclopedia of English Language Teaching*, (1st Ed.) pp.: 1-6. doi: 10.1002/9781118784235.eelt0325.
- Almazroue, H. (2009). The effectiveness of the alternative Assessment in the achievement of the concepts of educational research and increasing self-efficacy in research among postgraduate students. *Journal of the Union of Arab Universities*. 54 (December), 127-155.
- Almorshid, Yousef. (2014). Levels of reflective thinking among Al-Jouf University students: A Cross-sectional study. *Taiba University Journal of Educational Sciences*, 9/2: 163-184.
- Almqati, Y. (2015). *Investigating the difficulties of using science teachers' alternative assessment strategies at secondary schools in the area of Hail in Saudi Arabia from their point of view*. (Unpublished MA theses). Umm Al Qura University, KSA.
- Alnajim (2013). The impact of using the original assessment on the developing problem-solving skills and the attitudes towards fiqh course at the primary level. *Journal of college of education*. Ain Shams University –Egypt. 37/2: 231- 245.
- Alolyan, F. (2014). Trends of Mathematics teachers in the intermediate stage towards the use of alternative assessment in mathematics learning. *Risālat of Education and Psychology- Saudi Arabia*. 45: 53-76.
- Al-Omari, W. & Shehadeh, F. (2010). Degree of satisfaction of science teachers on the use of realistic evaluation methods in the teaching process. *Journal of the Faculty of Education*, Ain Shams University, Egypt, 34, 249-284.
- Al-Rashidi, F. (2018). The level of using reflective practices among secondary school Teachers in Baredah governorate. *Journal of Faculty for Basic Education in Educational and Human Sciences*, Babel University, 38: 284-294.
- Al-Rfou'a, M. (2017). The degree to which tenth-grade students possess reflective thinking skills and its relationship with their academic achievement. *Journal of the Faculty of Education, Al- Azhar University*, 174(1), 721-752.
- Al-Rifaa'i, A. Tawalbeh, H. Al-Qaood, I. (2012). The degree of using strategies of alternative evaluation by teachers of social studies in Irbid Governorate. *Journal of 'Um al Quraa for educational and psychological sciences*. 4/1: 369-408.
- Al-Sa'ydeh, N. M. (2016). Reflective thinking and its relationship with some demographic variables among gifted students in King Abdullah II Schools for Excellence in Jordan. *Studies in Educational Sciences*, Jordan, 43/4: 1747-1757.

- Al-Sharaa, N. & Zaza, H. (2013). Investigating elementary school teachers' assessment practices in Jordan: Towards a comprehensive and integrated model. *Journal of Educational and Psychological Sciences*, 14/2: 73-104.
- Alsuwaihiri, A. (2013). *Evaluation of tests and measurements of alternative assessment in mathematics for the upper grades in the primary stage in Makkah*. (Unpublished MA thesis). Umm Al-Qura University, KSA.
- Angel, G. (2015). *PISA 2015 Results in Focus*. *Andreas.Schleicher@oecd.org*. paper, see also [online]. [Accessed 16 March 2019]. Available from the World Wide Web: <http://www.oecd.org/pisa/pisa-2015-results-in-focus.pdf>.
- Ashkar, M. (2015). *The impact of employing alternative evaluation in development mathematical thinking for the fourth graders in Gaza*. (Unpublished Master Thesis), Faculty of Education, Islamic University, Gaza.
- Awawdeh, S. J. & Nasser, A. F. (2018). Does training in alternative assessment matter? The case of prospective and practicing mathematics teachers' attitudes toward alternative assessment and their beliefs about the nature of mathematics. *International Journal of Science and Mathematics Education*, 16/7: 1315-1335.
- Borg, S. (2003). Teacher cognition in language teaching: A review of research on what language teachers think, know, believe, and do. *Language Teaching*, 36: 81-109. <https://doi.org/10.1017/S0261444803001903>.
- Borko, H., Mayfield, V., Marion, S., Flexer, R. & umbo, K. (1997). Teachers' developing ideas and practices about mathematics performance assessment: Successes, stumbling blocks, and implications for professional development. *Teaching & Teacher Education*, 13/3: 259-278.
- Bramwell-Lalor S. & Rainford M. (2014), The effects of using concept mapping for improving advanced level Biology students' lower and Higher-order cognitive skills, *Int. J. Sci. Educ.*, 36/5: 839–864.
- Brown, J. B. & Hudson, T. (1998). The alternatives in language assessment. *TESOL Quarterly*, 32/4: 653–75.
- BuQhoos, K. (2017). Relationship between reflective thinking and teaching performance of student teachers specialized in science and mathematics. *International Journal for Research in Education*. 41/1: 38-65.
- Chicaiza, L. N. W. G. (2018). *Problem-based approach to develop the speaking skill in English*. *Magister dissertation*. Ambato, Ecuador.
- Diab, S. (2011). The impact of the use of various evaluation tools of seventh grade student achievement in science and attitudes towards learning. *Research Journal of Najah University*. 3/26: 33 - 48.
- Fook, C.Y and Sidhu, G. Kaur. (2014). Assessment practices in higher education in the United States. *Procedia- Social and Behavioural Sciences*, 123: 299-306. doi: 10.1016/j.sbspro.2014.01.1427.
- Gay, L. R. & Airasian, P. W. (2003). *Educational research: Competencies for analysis and application*. (7th Ed.), Prentice Hall. USA.
- Gay, L.R., Mills, G.E., & Airasian, P. (2009). *Educational research: Competencies for analysis and applications*. (9th Ed.). Upper Saddle River, NJ: Pearson, USA.
- Gozuyesil, E & Tanriseven, I. (2017). A Meta-analysis of the Effectiveness of Alternative Assessment Techniques. *Eurasian Journal of Educational Research* 70: 37-56. doi: <http://dx.doi.org/10.14689/ejer.2017.70.3>.
- Hamid, S. (2013). *The impact of using the alternative assessment methods in developing reflective thinking and mapping skills in geography of tenth grade students*. (Unpublished Master Thesis), Faculty of Education, Islamic University, Gaza. <http://hdl.handle.net/20.500.12358/22235>.
- Hassan, M. (2013). A reflective teaching-based program for developing teaching skills in accordance with quality standards and improving the teaching theory of Arabic language and Islamic pre-service teachers in Egypt and Saudi Arabia. *International Journal of Education*, 2/7: 652-682.

- Jangi, M. and Farjami, H. (2017). The contribution of collaborative assessment of classroom Participation to learners' sense of classroom community. *International Journal of Language Education and Applied Linguistics*, 31-46. Available online at <http://ijleal.ump.edu.my/>.
- Juhaina, A. S. & Nasser, A. F. (2018). Does training in alternative assessment matter? The case of prospective and practicing mathematics teachers' attitudes toward alternative assessment and their beliefs about the nature of mathematics. *International Journal of Science and Mathematics Education*, 16/7: 1315-1335.
- Martin, M. O., Mullis, I. V., Foy, P., & Hooper, M. (2016). *TIMSS 2015 International Results in Science*. Publisher: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College, Boston College, Chestnut Hill, MA 02467, USA.
- Napoli, A. R. & Raymond, L. A. (2004). How reliable are our assessment data? A comparison of the reliability of data produced in graded and un-graded conditions. *Research in Higher Education*, 48/8: 921-929.
- Nevo, D. (1995). *School-based Evaluation: A Dialogue for School Improvement*. Oxford: Elsevier/Pergamon.
- Obiedat, D., Kayed, A. & Adass, A. (2016). *Scientific research: understandable, tools and methods*. Jordan, Amman: Dar Alfikr Publishers and distributors.
- Odeh, K. (2015). *The impact of using authentic assessment on ninth grade students' achievement and their attitudes towards sciences in Nablus*. (Unpublished Master Thesis), An-Najah National University, Nablus, Palestine.
- Oliver, E. (2019). The triangle of effective education implemented for Theology. In *HTS Theological Studies*, 75/1: 1-8. a5234. <https://doi.org/10.4102/hts.v75i1.5234>
- Ostaz, M. H. (2011). The level of reflective thinking ability among teachers of science in the Fundamental School in Ghaza. *Journal of Human Sciences, Al-Azhar University*, 13/1b. PISA. (2015). See the link: <http://www.oecd.org/pisa/>
- Rodomanchenko, A. (2017). Roundtable discussion in language teaching: Assessing subject knowledge and language skills. *Journal of Language and Education*, 3/4: 44-51. <https://doi.org/10.17323/2411-7390-2017-3-4-44-51>.
- Ryan, A. (2014). The degree of reflective practices of mathematics teachers in Hebron Directorate of Education & its relationship with teaching self-efficacy. *Manara Journal*, AL-albait University, 20/1b: 143-173.
- Sandford, B. A & Hsu, C. C. (2013). Alternative assessment and portfolios: Review, reconsider, and revitalize. *International Journal of Social Science Studies*, 1/1: 215-221. doi:10.11114/ijsss.v1i1.69.
- Shaikh, O., Qudah, K. & Abd, E. (2007). *A Survey Study of Teachers', School Principals' and Parents' Perceptions of New Curricula and Textbooks*. National Report. National Center for Human Resources Development (NCHRD), Amman, Jordan.
- Taf, R. T., Dorf, Y. J., & Lazarowitz, R. (2000). A project-based alternative assessment system. *Studies in Educational Evaluation*, 26: 171-191.
- TIMSS. (2015). See the link: <http://www.timss.org/>
- Wikstrom, N. (2007). *Alternative assessment in the primary year of international baccalaureate education*. The Stockholm Institute of Education, Thesis 15ECTS.
- Zimbicki, D. (2007). *Examining the Effects of Alternative Assessment on Student Motivation and Self-Efficacy*. Doctorate thesis, School of Education: Walden University.