RES MILITARIS
REVUE EUROPEENNE D ETUDES EUROPEAN JOURNAL OF MILITARY STUDIES
Social Science Journal

Assessment of Jordanian Science Teachers Perceived Skill in Classroom Assessment

By

Ibrahim Mohammad Hussein Bani-Hamad
Universiti Malaysia Perlis

Amani Ali Elmetwaly
Universiti Malaysia Perlis

Junainor Hassan
Universiti Malaysia Perlis

Abstract

This study aims to investigate the students' learning assessment practices used by Jordanian science teachers. The sample of the study consists of (320) teachers. A questionnaire of (25) items are adapted on six dimensions, namely: Criterion Reference testing,Grading practices, Statistic applications, Assessment application,Essay items, and Objective items. Validity and reliability are established. Results of the study show that the mean of the scores for the six domains and the entire items are medium, the highest was on the Assessment application and the lowest on Statistical applications. Implications of these findings for policy makers and school managers are discussed.

Keywords: teacher perceived skill in classroom assessment; assessment practices; teachers self-perceived skills.

1. Introduction

Educational assessment is an essential component of the teaching profession. It is the process used in the classroom by the teacher to obtain information about students' performances on assessment tasks, using a variety of assessment methods, to determine the extent to which students are achieving the target instructional outcomes. In this regard, researchers suggest that a sound educational assessment requires a clear conception of all intended learning outcomes of the instruction and a variety of assessment procedures that are relevant to the instruction, adequate to sample student performance, and fair to everyone. This means teachers should competently be able to choose and develop assessment methods appropriate for instructional decisions; administer, score, and interpret results of externally produced and teacher-made assessment; use assessment results when making educational decisions; develop valid grading procedures; communicate assessment (Bueno & Figuerres, 2017).

Classroom assessment is an integral part of the teaching and learning process (Ashraf & Zolfaghari, 2015). Assessment has been affirmed as a vital component in professional careers of teachers. It enables them to improve their instructional practices and the learning progress of students (Gotch & French, 2014; Hussainetal.,2018). As a part of teaching learning, assessment becomes a procedure used by teachers to find out their students' level of knowledge and skills, learning outcomes, the strengths weaknesses, so that "they" can motivate them to improve their abilities. Additionally, the assessment makes teachers provide corrective feedback on what students are learning how much learning

material well students taught (Box et al., 2015). The aims of classroom assessment are not only explicitly intended to determine the weaknesses and success of the students, but also to figure out the ways of improving the quality of students learning (Ashraf & Zolfaghari, 2018).

Just like other Asian countries, Jordan has made tremendous efforts to make educational reforms, in 2003 the MoE launched a large education reform program, Education Reform for Knowledge Economy (ERfKE), aiming to enhance the education system quality and producing graduates with knowledge economy skills. These include problem-solving skills, analytical thinking skills, computer technology skills, communication skills, etc. (NAfKE report 2007). With this vision, the MoE soon launched the new curriculum reform and development for all grades (in multiple phases) to prepare students for life-long learning and mastery of the new skills. In 2006 it implemented the first phase of the new textbooks and teaching methods (e.g. promoting student centered teaching methodology, multi-facet ways of learning, collaborative learning, etc.) and new assessment tools for Grades 1, 4, 8, and 10. The curricula change for other grades were followed in a systematic process (Ababneh et al., 2014).

Teachers are the primary assessors inside the classroom who usually use the tools of assessment in order to get information about teaching and learning, in addition to monitoring students' progress in achieving learning objectives. Their role in this side exceeds just assisting students to develop self-monitoring and self-assessment skills and strategies, teachers work on driving their students to be involved in setting learning outcomes, developing plans, and using alternative strategies in monitoring their own achievement. Furthermore, teachers watch students' learning and progress by observing them regularly and systematically in classrooms by interacting with them while teaching. Teachers collect data by observing and assessing students' interaction, performance, and projects or work samples, and making judgments based on these observations demonstrating that assessment is an essential part of the learning process. The effective alternative strategies "they" model in the classroom lead to involve the students in the procedures of developing assessment such as preparing rubrics and checklists(Abed & Abu Awwad, 2016) .

Teacher's assessment knowledge and the teachers' assessment competence have a direct effect on teaching and learning process in the classroom. This is because teachers need to use assessment information to make informed decisions about students' learning and communicate assessment results effectively. Therefore, teacher competency and knowledge of classroom assessment is directly related to effective student teaching (Widiastuti et al., 2021). Zhang and Burry-Stock (2003) argued that teachers' perceived skill in classroom assessment practices reflects their perceptions on their skill in conducting classroom assessment practices. Zhang and Burry-Stock explain why teachers may rate their assessment skills as good even if they are found to be incompetent to conduct some assessment practices. When asked about assessment training they received and if such training benefited their classroom practices, teachers generally indicate that assessment training they receive did not adequately prepare them for their classroom assessment practices (Koloi-Keaikitse, 2012; Mertler, 2009). As educators we are all aware of implications of limited teacher assessment training on teaching and learning. Insufficient teacher training may impact on skill-based assessment needed for sustainable development. It is therefore essential to develop teachers' assessment competencies and skills to improve their classroom assessment practices to cope with the ever changing twenty-first-century educational needs. Chester & Chester and Quilter

(1998) as cited in Susuwele-Banda (2005) argued that it is important to study teacher perceptions of assessment as this can inform educators how different forms of assessments are used or misused and what can be done to improve teacher classroom assessment practices. DarlingHammond et al. (2012) as cited in Koloi-Keaikitse (2016) supports the need for teacher competencies in assessment because if teachers feel prepared when they enter teaching they are most likely to have better sense of teaching efficacy which can ultimately improve their motivation to teach. Unfortunately, teachers lack adequate knowledge and competence regarding classroom assessment. Mostly, teachers were found not to be good judges of the quality of their own assessment activities as well as their students' abilities (Mellati & Khademi,2018 ; Clark-Gareca, 2016). This study was therefore meant to assess teachers' response pattern on a self-perceived skills in classroom assessment practices. Teachers' assessment practices are an essential element for addressing students' learning needs, and they can ultimately improve the education system and accountability. Understanding teachers' assessment practices serve as a way of finding out if teachers adopt or use quality assessment methods that can address the learning needs of students (McMillan, 2001).

Purpose of the study

The study identified items that provide the most information about teachers' perceived skills in classroom assessment practices. In turn, by identifying which skills teachers are most and least confident about, it is hoped to provide educational administrators, policy-makers, and teacher educators with useful information for the planning and conducting of assessment training for teachers. The main research question is; which classroom assessment practices do Jordanian science teachers perceive themselves more skilled?

Method

Study population and sampling

Population and Sample the present study was carried out in middle schools in Jordan, specifically in six directorates located at the north of Jordan: Bani Obeid, Northern Jordan Valley, Ramtha, Qasbah Irbid, North Mazar and Koura. The research population sample selected for the present study included include ((1700) science teachers(711) 44.3% male and (894) 55.7% female who teach Grades (4-8) enrolled in the Ministry of education public schools . according to Krejcie and Morgan (1970), it is appropriate to select a minimum sample of 313 steachers from the entire research population. The sample comprised of 320 science teachers (142) 44.3% male and (178) 55.7% female. This ensured that the number represented the whole population and schools.

Instrument

The Classroom Assessment Practices and Skills (CAPS) Questionnaire adapted by Koloi (2017) from Zhang and Burry (2003) was used as the data collection instrument to assess teachers' perceived skills in classroom assessment practices.teachers were asked to indicate their perceived classroom assessment skills, such as constructing objective items, conducting item analysis for teacher made tests, using non- achievement factors such as students' effort, motivation, and improvement when awarding grades to students, and interpreting tests results. It designed to measure teachers' self-perceived from 1 (not at all skilled) to 5 (very skilled), Items were scored so that higher numbers indicated higher

perceived skill in classroom assessment practices. Classroom Assessment Practices and Skills consisting of (29) items and divided into six dimensions as follows; (4) items for Criterion Reference testing,(6) for Grading practices,(6) for Statistic applications,(9) for Assessment application,(2) for Essay items ,and (2) for Objective items. Cronbach's Alphas for the total Questionnaire was (α = .95) and its domains were as follows: Criterion Referenced Testing (α =.88), Grading Practices (α =.82), Statistical Applications (α =.84), Assessment Applications (α =.88), Essay Items (α =.75), and Objective Items (α =.86).

2.2.1 Validity

In order to ensure that items were content and context relevant for teachers in Jordan based on the language and curriculum, Questionnaire validity was verified by (6) arbitrators from supervisors and faculty members. Appropriate. Based on that revision process all items were found to be relevant, and the number of the items remained (29).

2.2.2 Reliability

To investigate reliability, the questionnaire was applied on a sample of (50) teachers out of the participants. Six dimensions (Criterion Reference testing, Grading practices, Statistic applications, Assessment application, Essay items, and Objective items measured the internal consistency of the Cronbach's alpha for the teacher self-perceived skills and practices. The first run of the data showed that the reliability value was more than the suggested value of 0.60 for all teacher self-perceived skills and practices dimensions. However, the first dimension (Criterion Reference testing) consists of 4 items, and all achieved more than 0.30 values for items correlation the second dimension (Grading practices) consists of 6 items, 4 of which achieved more than 0.30 values for items correlation. The third dimension (Statistic applications) consists of 6 items, and all achieved the suggested value of .30 for items correlation.the fourth dimension (Assessment application) consists of 9 items, 7 of which achieved the suggested value of .30 for items correlation. The fifth dimension (Essay items) consists of 2 items and all achieved the suggested value of .30 for items correlation. Finally, dimension six (Objective items) consists of 2 items, and all achieved the suggested value 0.30 of items correlation.so,the total items of the questionnaire are (25). Internal Consistency was calculated through Cronbach's Alpha Coefficient, it was 0.86 for the whole questionnaire. Reliability of the sub dimensions ranged between .760 and .905 Table 2 showed these results.

Table 1: *Reliability Analyses for Teacher assessment skills in classroom assessment Instrument*

dimension	Number of items	Alpha co.
Criterion Reference testing	4	.806
Grading practices	4	.760
Statistic applications	6	.802
Assessment application	7	.807
Essay items	2	.876
Objective items	2	.905

2.3. Data analysis

Means and standard deviations were calculated for each item and the overall score of the questionnaire using spss version 28. Items were ranked according to their means in a

descending order as shown in table 3. Means were categorized into three levels: low (1 to 2.33), medium (2.34 to 3.66), and high (3.67-5).

2.4. Results

Means and standard deviations for assessment skills items and level of each

Item	mean	Sta.deviation	level
Criterion Reference testing			
Assessing specific course objectives	4.1719	.82581	High
Ensure test covers the material taught	2.9844	1.08411	Medium
Fairly assign grades to all students	2.5938	1.15450	Medium
Align test items with instructional obj	2.6563	1.18532	Medium
Grading practices			
Including student effort when grading	2.7563	1.18373	Medium
Develop systematic grading procedure	3.1594	1.17560	Medium
Use table of specification for test plan	4.0031	1.08999	High
Develop rubrics (marking keys)	3.9063	1.06407	High
Statistic applications			
Explain standardized exam scores	3.1969	1.31378	Medium
Calculate central tendency	3.1812	1.25881	Medium
Conduct item analysis	3.1531	1.24882	Medium
Revise items based on item analysis	1.8938	.82361	low
Use peer assessments	3.3875	1.22915	Medium
Calculate variability	3.2438	1.22792	Medium
Assessment application			
Write items for higher cognitive level	4.2125	.88092	High
Assess individual class participation	3.2406	1.27978	Medium
Assess problem-solving skills	3.2656	1.24495	Medium
Use portfolio assessment	3.1000	1.28263	Medium
Use assessment results for decisions	3.0125	1.31777	Medium
Determine why students make mistakes	3.2688	1.21754	Medium
Use assessment results to plan teaching	3.3000	1.23329	Medium
Essay items			
Construct essay items	3.3187	1.21574	Medium
Consistently grade essay question	3.1781	1.29968	Medium
Objective items			
Construct multiple choice items	3.2719	1.34996	Medium
Construct true or false items	3.1000	1.31163	Medium
first domain: Criterion Reference testing	3.1016	.70505	Medium
second domain: Grading practices	3.4562	.66868	Medium
third domain: Statistic applications	3.0094	.79109	Medium
fourth domain: Assessment application	3.3429	.87942	Medium
fifth domain: Essay items	3.2484	1.15967	Medium
sixth domain: Objective items	3.1859	1.21921	Medium
over all	3.185	.7068	Medium

Results in Table 3 revealed the following: According to the first domain (Criterion Reference testing), the item “Assessing specific course objectives ” had the highest mean among the items in the domain while the item “ Fairly assign grades to all students ” had the lowest mean. There is one item that got high level of practice, and (3) items got medium level

of practice, but there weren't any item that got low level of practice. According to the second domain (Grading practices), the item "Use table of specification for test plan" had the highest mean among the items in the domain while the item "Including student effort when grading" had the lowest mean of the items. There were (2) items that only got high level of practice, and (2) items got medium level of practice, but there weren't any item that got low level of practice. According to the third domain (Statistic applications), the item "Use peer assessments " had the highest mean among the items in the domain while the item "Revise items based on item analysis ' questions" had the lowest mean of the items. There were (5) items that got medium level of practice, and one item got low level of practice, but there weren't any item that got high level of practice. According to the fourth domain (Assessment application), the item "Write items for higher cognitive level" had the highest mean among the items in the domain while the item "Use assessment results for decisions " had the lowest mean of the items. There is one item that got high level of practice, and (6) items got medium level of practice, but there weren't any item that got low level of practice. According to the fifth domain (Essay items), the item "Construct essay items" had the highest mean among the items in the domain while the item "Consistently grade essay question" had the lowest mean. There were (2) items got medium level of practice, but there weren't any item that got low or high level of practice. According to the sixth domain (Objective items), the item " Construct multiple choice items" had the highest mean among the items in the domain while the item "had the lowest mean. There were (2) items got medium level of practice, but there weren't any item that got low or high level of practice. Results in Table 3 revealed also that the mean of the (Grading practices) domain was the highest (3.4562), then the mean of the (Assessment application) domain was (3.3429), then the mean of the (Essay items) domain was(3.2484), then the mean of the (Objective items) domain was (3.1859) , then the mean of (Criterion Reference testing) domain was(3.1016), while the mean of the (Statistic applications) domain was (3.0094) the lowest. It is noted from Table 3. That the overall mean of the assessment practices equals (3.185), and the standard deviation for these practices is (.7068). Table (3) shows that the average of the assessment practices in the (Grading practices) domain ranges between (2.7563- 4.0031) with a total mean of (3.4562), and the standard deviation for such practices ranges between (1.06407 -1.18373). The table also shows that the average of the assessment practices of (Essay items) ranges between (3.1781 -3.3187) with a total mean of (3.2484), and the standard deviation for such practices ranges between (1.21574-1.29968). The table also shows that the average of the assessment practices of (Objective items) ranges between (3.1000- 3.2719) with a total mean of (3.1859), and the standard deviation for such practices ranges between (1.31163 - 1.34996)). The table also shows that the average of the assessment practices of (Criterion Reference testing) ranges between (2.5938-4.1719) with a total mean of (3.1016), and the standard deviation for such practices ranges between (.82581-1.18532). The table also shows that the average of the assessment practices of science teachers in regard to the (Statistic applications) domain ranges between (1.8938-3.3875) with a total mean of (3.0094), and the standard deviation for such practices ranges between (.82361-1.31378).It is also noted from Table 3 that the most alternative practice in the Assessment application domain is the Write items for higher cognitive level with an average of (4.2125). Also, table (3) shows that the least assessment practice of teachers in the Statistic applications domain is revise items based on item analysis with an average of (1.8938).

4. Discussion and conclusion

The study purpose was to provide a rich description of Jordanian science teachers' for grades (4-8) skill in classroom assessment. Given the value for classroom assessment

practices and the need to have a clear understanding of teachers perceived skills in classroom assessment practices (Cavanagh, Waldrip, Romanoski, Dorman, & Fisher, 2005; Marriot & Lau, 2008; Nenty et al., 2007; Rowntree, 1987; Zhang & Burry-Stock, 2003) as cited in Koloi (2017) and the relative scarcity in this area of research in Jordanian context, the study adds to the understanding of teachers perceived skills in classroom assessment practices.

Teachers continue to be important for bringing change and preparing students for future endeavors. It is therefore imperative to understand their teaching practices particularly how they assess and evaluate student learning outcomes (Reynolds et al., 2009; McMillan, 2020). Gathering information that can highlight the level of teachers' classroom assessment competences in conducting classroom assessments is vital to determine their capabilities and inadequacies. Such information can be used by institutions that conduct teachers' education and professional development to develop teachers' assessment skills that teachers perceive to possess as well as those for which they feel less competent.

The study purpose was to provide a rich description of Jordanian science teachers' for grades (4-8) skill in classroom assessment. Given the value for classroom assessment practices and the need to have a clear understanding of teachers perceived skills in classroom assessment practices (Cavanagh, Waldrip, Romanoski, Dorman, & Fisher, 2005; Marriot & Lau, 2008; Nenty et al., 2007; Rowntree, 1987; Zhang & Burry-Stock, 2003) as cited in Koloi (2017) and the relative scarcity in this area of research in Jordanian context, the study adds to the understanding of teachers perceived skills in classroom assessment practices. Understanding teachers' perceptions about their perceived skills in classroom assessment practices is very important as it can open avenues informing policy and practice for addressing the needs that teachers have as they wrestle with their day-to-day classroom assessment practices (koloi, 2017).

Teachers continue to be important for bringing change and preparing students for future endeavors. It is therefore imperative to understand their teaching practices particularly how they assess and evaluate student learning outcomes (Reynolds et al., 2009; McMillan, 2020). Gathering information that can highlight the level of teachers' classroom assessment competences in conducting classroom assessments is vital to determine their capabilities and inadequacies. Such information can be used by institutions that conduct teachers' education and professional development to develop teachers' assessment skills that teachers perceive to possess as well as those for which they feel less competent.

Overall results of this study revealed that jordainian science teachers are most skilled in the (Grading practices) domain such as "Use table of specification for test plan" and "Develop rubrics marking keys" followed by (Assessment application) domain such as Write items for higher cognitive levelwhile the (Statistic applications) such as Revise items based on item analysis was the least. Additionally the results showed that an item that asked teachers about their perceived skill in Statistical applications proved to be the most difficult for respondents to endorse, an indication that most of the teachers were less skilled in test items analysis. Further research to establish why teachers felt least competent in statistical item analysis is highly recommended.

These findings have major implications for teacher educators and school managers. For teacher educators these results highlight classroom assessment areas that they may need to focus on as they teach assessment courses. Assessment entails a broad spectrum of activities that includes test itemes analysis in order to revise items based on item analysis and collection of information about the students learning for decision-making. The

responsibility of teachers is to collect information through various assessment methods that can be used to make informed decisions about students' learning progress. The question is: are teachers competent enough to use or apply tests with good items to assesst their students for making students' learning decisions? From these results it was very clear that teachers are less confident in good items to assesst their students for making students' learning decisions.

It is clear that items that assessed jordainian science teachers' perceived skills about classroom assessment skills showed that they are less skilled in the statistical applications. This finding is important because it shows that if the Ministry of Education want to know assessment areas that teachers may need to be trained on, it may concentrate on statistical application more than other classroom assessment skills activites and workshops. These results generally imply the need for teacher educators or assessment professional development specialists to focus their attention on assessment training on skills teachers need most and those they have less perceived skills on.

5. Recommendations

The study presented the following recommendations:

- Further research to establish why teachers felt least competent in statistical item analysis .
- more in-survece assessment training for teachers in statistical application by the Ministry of Education to becomes more confident in items analysis.

References

Ababneh, A., Ababneh, I., Lebdih, K. A., & Tweissi, A. (2014). Mapping of Student

Abed, E. R., & Abu Awwad, F. M. (2016). Students' Learning Assessment Practices Used by Jordanian Teachers of Mathematics for Grades (1-6). International Education Studies, 9(1), 63-78.

And Ideas, 93(2), 85-92.

Ashraf, H., & Zolfaghari, S. (2018). EFL Teachers' Assessment Literacy and Their Reflective Teaching. International Journal of Instruction, 11(1), 425-436

Assessments in Jordan.

Clark-Gareca, B. (2016). Classroom assessment and English Language Learners: Teachers' accommodations implementation on routine math and science tests. Teaching and

Development, 20(2), 107-123.

Koloi-Keaikitse, S. (2012). Classroom assessment practices: A survey of Botswana primary and secondary school teachers. Ball State University.

Koloi-Keaikitse, S. (2016). Assessment training: A precondition for teachers' competencies and use of classroom assessment practices. International Journal of Training and

Koloi-Keaikitse, S. (2017). Assessment of teacher perceived skill in classroom assessment practices using IRT Models. Cogent Education, 4(1), 1281202.

McMillan, J. H., & Moore, S. (2020). Better being wrong (sometimes): classroom assessment that enhances student learning and motivation. : A Journal of Educational Strategies, Issues

McMillan, J. M. (2001). Secondary teachers' classroom assessment and grading practices. Educational Measurement: Issues and Practices, 20, 20–32

McMillan, J. M. (2008). Assessment essentials for student-based education (2nd ed.). Thousand Oaks, CA: Crown Press.

Mellati, M., & Khademi, M. (2018). Exploring teachers' assessment literacy: Impact on learners' writing achievements and implications for teacher development. Australian Journal of

Mertler, C. A. (2009). Action research: Teachers as researchers in the classroom. Sage.

Redondo, S. C., Bueno, D. C., & Figuerres, N. R. (2017). Analysis of classroom assessment skills and practices across levels in a Catholic educational institution. CC: The Journal, 13.

Susuwele-Banda, W. J. (2005). Classroom assessment in Malawi: Teachers' perceptions and practices in mathematics. Virginia Polytechnic Institute and State University.

Teacher Education (Online), 43(6), 1-18.

Teacher Education, 54, 139-148.

Widiastuti, I. A. M. S., Mantra, I. B. N., Sukoco, H., & Santosa, M. H. (2021). Online assessment strategies to enhance students' competence and their implementational challenges.

Widiastuti, I. A. M. S., Mantra, I. B. N., Sukoco, H., & Santosa, M. H. (2021). Online assessment strategies to enhance students' competence and their implementational challenges.

Zhang, Z. R., & Burry-Stock, J. A. (2003). Classroom assessment practices and teachers 'self- perceived assessment skills. Applied Measurement in Education, 16(4), 323-342.