Effectiveness of the Conceptual Change Approach in Improving Teachers' Conceptions of Assessment at the University of Bahrain

Numan M. Al-Musawi
University of Bahrain- Kingdom of Bahrain
nalmosawi@hotmail.com

Received: 10 Sept. 2015
Revised: 13 Feb. 2015, Accepted: 09 March. 2015
Published online: 1 (April) 2016

International Journal of Research in Education and Psychology
ISSN 2210 -1780
Int. J. Res. Edu. Psy. 4, No. 1 (Apr. 2016)

Effectiveness of the Conceptual Change Approach in Improving Teachers' Conceptions of Assessment at the University of Bahrain

Numan M. Al-Musawi
University of Bahrain - Kingdom of Bahrain

Abstract

The purpose of this study was to examine the effectiveness of an innovative approach to teacher development, the conceptual change approach, that attempts to change teachers' perspectives on assessment. Participants were twelve in-service teachers involved in a staff development program for teachers at the University of Bahrain. The evaluation investigated the program at three levels: The impact of the program on the participants' conceptions of assessment, the resultant impact on their assessment practices, and the consequential effects on their students' gains in achievement. The mixed methods design was used in theory-driven evaluations under the conceptual framework of Program Theory (Ho, 2000). Results showed that the program brought about detectable conceptual change in four teachers whose students received higher grades in the following year, while none of the students of those teachers who did not change their conceptions about assessment showed similar gains in achievement. Implications for program evaluation theory and assessment practice in schools and higher education institutions are discussed.

Keywords: Program theory, conceptual change approach, professional development programs, conceptions and practices of assessment, school teachers.

فاعلية طريقة تغيير المفاهيم في تحسين تصورات المعلمين للتقييم في جامعة البحرين

نعمان محمد الموسوي

جامعة البحرين – مملكة البحرين

الملخص

تهدف الدراسة الحالية الى تحديد درجة فاعلية مدخل "تغيير المفاهيم "، وهو مدخل جديد لتطوير المعلّم يسعى إلى تغيير آراء المعلّمين وتصوّراتهم بشأن التقويم. اشترك في الدراسة الحالية اثنا عشر معلماً أثناء الخدمة، وهؤلاء المعلّمون مشاركون في برنامج للتنمية المهنية للمعلّمين بجامعة البحرين، حيث تم تقويم البرنامج على ثلاثة مستويات، وهي: أثر البرنامج في تصوّرات المعلّمين حول التقويم، وأثر ذلك على ممارساتهم التقويمية الصفيّة، والأثر المترتّب من جرّاء ذلك على مستوى تحصيل طلبتهم. استُخدم في الدراسة تصميم الطرق المختلطة (الكميّة والنوعيّة) في دراسة أساليب التقويم القائمة على أساس النظرية، وذلك على خلفيّة الإطار المفاهيمي لنظرية البرنامج (Ho, 2000). وبيّنت نتائج الدراسة أن البرنامج أحدثَ تغييراً قابلاً للملاحظة في المفاهيم المتصلة بالتقويم لدى أربعة معلّمين فقط حصل طلبتهم على تقديرات مرتفعة في السنة اللاحقة، أما أولئك المعلّمون الذين لم يغيّروا مفاهيمهم المتعلقة بالتقويم فإن أياً من طلبتهم لم يتمكّن من تحسين مستوى تحصيله بالمثل. وعلى هذا الأساس، تم مناقشة تطبيقات تائج الدراسة على صعيد نظرية تقويم البرنامج، وكذلك الممارسات التقويمية في المدارس مؤسّسات التعليم العالي.

الكلمات المفتاحية : نظرية البرنامج، مدخل تغيير المفاهيم، برامج التنمية المهنية للمعلمين، تصوّرات التقويم، ممارسات التقويم الصفّي، معلّم المدرسة.

Effectiveness of the Conceptual Change Approach in Improving Teachers' Conceptions of Assessment at the University of Bahrain

Numan M. Al-Musawi

University of Bahrain - Kingdom of Bahrain

INTRODUCTION

The widespread demands from both the governments and the consumers for greater accountability, in conjunction with the quality improvement movement, have resulted in a need for valid, reliable and comparable performance data on the quality of teaching in higher education institutions. This, in turn, has activated the search for new assessment approaches and also for innovative models for the professional development of teachers (Madina, Baker, Chow, Delacruz, & Griffin, 2015).

One of the main components of effective teaching and learning at the university is the assessment of students (March, 1987). Assessment is a process that provides information about the thinking, achievements or progress of students (McMillan, 2007). If this activity is not well planned and effectively implemented, it fails to promote student learning and achieve the desired goals of teaching and learning. Conventionally, assessment practices in the classroom might take the form of quizzes, unit tests, and formal assignments, standardized diagnostic assessments such as teacher-made pre-assessments or formative assessments. The teacher-directed assessment practices also include the use of techniques such as personal whiteboards, student conferences, and even watching pupil's facial expressions. This spectrum of classroom-based, assessment practices represents the means by which a teacher candidate gathers relevant and dependable information about the degree of student learning so that he or she can then communicate with others, draw inferences, make proper decisions, and take instructional actions (Hargreaves, 2007; Gordon, 2008).

In the contemporary education context, however, assessment has taken on new roles beyond progressing student learning. The passage of No Child Left Behind in the United States in 2001 established federal expectations for the assessment of student learning by each of the states as an accountability measure. Recently, states such as Virginia have enacted standards that require the demonstration of student progress as a significant component of teachers' evaluations (Virginia Board of Education, 2011). These movements have placed increasing importance on the role of assessment for purposes of evaluation - that is, the measuring of student learning in order to render judgments of the effectiveness or value of instructional efforts.

Assessment is an essential component of a coherent educational experience. It is an integral component of institutional effort to improve the quality of teaching and learning at all levels of education, including higher education. The two, above-cited conceptions of teaching necessitate two conceptions about assessment: assessment of learning and assessment for learning (Lorna, 2003). The assessment of learning, referred to as "Summative Assessment", is used for the purpose of grading, reporting and making judgments about students' progress in learning, whereas the assessment for learning "Formative Assessment" has the central

purpose of improving learning and teaching for both students and teachers through efficient assessment and feedback. This prospective requires us to focus on how we use assessments for student success and not on how we assess student achievement in traditional ways (Stiggins, 2001). While the assessment for learning conception is associated with the teacher's efforts to make learning explicit and promote learning autonomy, the assessment of learning conception is associated with performance orientation and the teach-to-the-test methods (Warren & Nibset, 1999; Lorente-Catalån & Kirk, 2015).

Numerous researchers (e.g., Brown, 2004; Brown, 2009; Crooks, 1988; Torrance & Pryor, 1998) have suggested that five major purposes for assessment of students exist: improvement of teaching and learning, making students accountable for learning through certificates, accountability of schools and teachers, and assessment is basically irrelevant to the work and life of teacher and students.

The major premise of the improvement conception is that assessment improves students' own learning and the quality of teaching (Crooks, 1988; Black & Wiliam, 1998). This improvement has two important pillars; (a) assessment must describe or diagnose the nature of student performance and (b) the information must be a valid reliable and accurate description of student performance. In this view, a range of techniques, including informal teacher-based intuitive judgment as well as formal tool of assessment, identify the content and processes of student learning,

A second conception of assessment is that assessment can be used to account for a teacher's or a school's use of society's resources (Butterfield, Williams, & Marr, 1999; Guthrie, 2002; Smith, Heinecke, & Noble, 1999) and imposes consequences for schools or teachers for reaching or not reaching required standards. Indeed, two rationales for this conception exist; one emphasizes demonstrating publicly that schools and teachers deliver quality instruction (Hershberg, 2002), and the second and the second emphasizes improving the quality of instruction (Linn, 2000).

The premise of the third conception of assessment is that students are held individually accountable for their learning through assessment. This is seen in checking off student performance against criteria, and placing students into groups based on performance. The idea of the fourth conception of assessment is that the midterm or final exam or any kind of achievement test should mainly measure the extent to which students memorized the learned content. The premise of the final conception is that assessment as a formal, organized process of evaluating student performance has no legitimate place within teaching. Teachers' knowledge of students based on their understanding of curriculum and pedagogy preclude the need to carry out any kind of assessment beyond the intuitive in-the-head process that occurs automatically as the teachers interact with their students (Airasian, 1997).

Different researchers in the field of assessment and evaluation sought to change the teachers' conceptions and perceptions about assessment using approaches, such as conceptual change approach (Ho, 2000), or empowerment evaluation (Charoenchai, Phuseeorn, & Phengsawat, 2015), or by using Web tools, through the integration of technology into education for instructional and assessment purposes (Cirit, 2015)

While the empowerment under the guise of evaluation approach is grounded on a notion that the "external evaluator's efforts to empower a group to conduct its own evaluations are advanced as external or independent evaluations" (Stufflebeam & Shinkfield, 2007, p. 154), the conceptual change approach to staff development is based on an idea that if we wish instructors to adopt student-centered approaches to teaching and students to adopt meaningful approaches to learning, it is important to "direct staff development efforts towards changing teacher conceptions of teaching to emphasize the facilitation of student learning" (Ho, 2000, p. 31).

The most influential theory in conceptual change in teaching, namely in science, is the conceptual change theory (Posner, Strike, & Hewson, 1982) that attempted to tackle the

following question: 'How do learners make a transition from a conception to a successor conception?' This theory has roots in the philosophy of science and is based on the assumption that conceptual change in learning is analogous to the model of development in scientific beliefs among the community of scientists.

The conceptual change theory has two parts: the 'conceptual ecology', i.e., the conceptual context of the learner in which learning takes place, e.g., the learner's current conceptions and misconceptions; and the set of conditions required for conceptual change to occur, e.g., there must be dissatisfaction with current conceptions, and a new conception must be intelligible, plausible, and fruitful.

The central idea of Posner, et al.'s theory rests with the learner being confronted with a conceptual conflict: for instance, the inability of existing conceptions about a scientific concept to explain the observed phenomena. This confrontation will initiate a learning process that involves replacing the original conception with a scientifically accepted one, provided that the new conception satisfies the other conditions. While this theory is more concerned about the requisites for change to really occur, i.e., 'what makes a change viable?', the concept of 'theories-of-action' (Argyris & SchÖn, 1974), which refers to the underlying rationale that someone holds for a particular action, poses a confrontation in a form of different types of dilemmas that could create the necessary tension to provide the foremost and essential ignition key for a change to occur, identifying three major dilemmas: Dilemmas of incongruity between espoused theory and theory-in-use; Dilemmas of internal inconsistency between beliefs and values within the theory-in-use and Dilemmas of effectiveness that arise when it becomes increasingly difficult to achieve the intended goals.

On the basis of the theories of conceptual change that encompass varied and comprehensive perspectives about change, Ho (2000) managed to develop a theoretical framework for a conceptual change program, which included the following elements: the self-awareness process; the confrontation process; exposure to better, alternative conceptions; and the commitment building and refreezing. This model for conceptual change program, which contributed to the design and implementation of a four three-hour session staff development short course at the Hong Kong Polytechnic University aimed to produce changes in the conceptions of teaching of participants towards one that would be more conducive to quality student learning, will be used in this study.

To evaluate the effect of the conceptual change program, we referred to the conceptual framework of Program Theory, developed by Chen (2005). Program theory is defined as a set of explicit and implicit assumptions held by stakeholders about what actions are required to solve a social problem and why the problem will respond to these actions (Chen, 2005). Program theory consists of two models: an action model and a change model. The action model represents a systematic plan for arranging staff, resources, setting and support organizations in order to reach target populations and provide intervention services. The change model entails the following components: intervention, which refers to a set of program activities that focus on changing the determinants and outcomes; determinants, the mechanisms that mediate between intervention and outcomes; and outcomes, the anticipated effects of the Program.

In theory driven-evaluations, the need for program theory clarification and holistic assessment usually requires the use of mixed methods (Chen, 2006). Theory driven-evaluations require two primary tasks: (a) facilitating stakeholders in clarifying or developing their program theory, and (b) empirically assessing program theory. The wide scope of theory-driven evaluations involves the sequential combination of these two tasks, and the program theory shows the opportunities for using various strategies for applying mixed-methods, such as quantitative and qualitative methods.

At the University of Bahrain (UOB), assessment is considered as the foundation for institutional effectiveness and improvement. The overall aim is to ensure that the University and its academic programs meet and strive beyond

minimum standards. Furthermore, planning and assessment are also important elements in the process of international accreditation for most UOB academic programs. It is important to note here that the term "programs" is used here to refer to "all degrees offered by the University of Bahrain and its colleges and supported by the various services and administrative units" (Mohieldin, Al-Ammal, & Al-Burshaid, 2010, p. 28).

At the program level, each academic program at the University of Bahrain is expected to develop and implement a plan to assess student learning outcomes. To assess student learning in any academic or professional development program, the following five steps are taken: (1) identify mission, educational objectives, and intended outcomes; (2) define certain and clear criteria for success; (3) evaluate performance against criteria or indicators; (4) analyze assessment results, and (5) seek improvement through actions.

For example, assessment procedures in the Post Graduate Diploma in Education (PGDE), a two-year professional development program for in-service teachers enrolled in the Bahrain Teachers College (BTC), include examinations, tests and quizzes, research, oral and written reports, specific teaching and learning aids, student assessment rubrics and an electronic professional portfolio. Also, video-taped micro-teaching sessions, in-class student teaching, and demonstrations of best practice are important parts of ongoing assessment of in-service teachers (BTC, 2009, p. 61).

As for the professional development of in-service teachers, the traditional teacher development activities are largely concerned with teaching skills and methods, for example how to conduct lessons, how to use questioning and observation techniques, how to apply methods of assessment of learning in the classroom context, etc. Many teacher development programs, as pointed out by Ho (1998) and Ramsden (1992), work on an assumption that providing learners with prescribed skills and teaching recipes will produce better teachers, that participants will accept, acquire and adopt the skills introduced to them. The experiences of many teacher developers, however, have suggested that participants, in many cases, question the feasibility of the new methods presented, defend the methods that they use or apply new methods mechanically and are unable to extend the ideas into other situations (Carnell, 2007). Hence, a call from the responsible bodies for new programs of teacher development has gained wide acceptance among educators.

As a result, the conceptual change approach to teacher development has emerged as a promising means of achieving genuine development in school teachers (Biggs, 1989; Bowden, 1989; Gibbs, 1995). Staff developers have begun to spell out explicitly that educational development is itself a learning process for prospective teachers, and that effective development programs need to bring about conceptual changes. If this would take place in the field, teachers should have an adequate conception of assessment that facilitates the maximum level of student learning. Given that assessment, as a whole, used to discover what a student understands or does not understand, can be a powerful tool in targeting instruction so as to move student learning forward, it should be viewed by prospective teachers as a way to accurately diagnose the nature of student performance and to improve students' own learning and the quality of teaching (Black & Wiliam, 1998) rather than just a mathematical method by which students are graded against typical performance indicators adopted by the relevant higher education institution. In other words, the teacher's awareness of the goal and purposes of assessment is the vehicle by which teacher's assessment practices are directed at correct diagnoses and improvement of student achievement and the enhancement of quality of teaching.

Since the staff development programs are meant to instigate changes in the conceptual understanding of assessment process and activities that should enhance the quality of teaching and student learning, the principal aim of the current study, then, is to contribute to the development of the conceptual change approach to staff development by using the conceptual framework of Program Theory (Ho, Watkins, & Kelly, 2001) to assess the impact of a staff

development program for in-service teachers at Bahrain University on their assessment's conceptions and practices. Given that the traditional assessment conceptions prevail among teachers involved in the professional development programs at the University of Bahrain (Al-Musawi, 2003, p. 44), this study seeks to measure the impact of the conceptual change approach on the in-service teachers' espoused conceptions and daily classroom practices of student learning assessment.

THE PROBLEM STATEMENT

In the last two decades of educational reforms, large structural changes have been initiated in schooling and higher education sector in Bahrain (Higher Education Policy Institute for the Education Reform Abroad, 2010). The present Ministry of Education (MoE) of the Kingdom of Bahrain is a policy only body; while other governing and statutory bodies deal with important functions devolved from the MoE; specifically, the Education Reform Board which oversees the Bahrain's Higher Education Agenda (improvement strategy), the provider institutions (e.g., Bahrain University), regulators (such as the Quality Assurance Authority for Education and Training that monitors and provides improvement guidance for schools and tertiary level institutions), or policy makers (such as Higher Education Council).

Within the framework of the whole educational reform that aims to improve the processes and the outcomes of education in Bahrain (Quality Assurance Authority for Education and Training, 2009; 2012), the University of Bahrain, in collaboration with the Ministry of Education in Bahrain, seeks to enhance teaching and leadership skills of in-service teachers and school administrators as well through various programmes of professional development, including the Post Graduate Diploma in Education (PGDE). Students who are enrolled in the PGDE program are required to successfully pass 11 PD courses or an equivalent of 33 credit hours within two years of study in the evening period at University. The PGDE program in the Bahrain Teachers College at Bahrain University (2009), for example, targets university graduates with bachelor's degree and engages them in a "range of teaching-learning strategies such as lecture, tutorial, group work, role play, interactive communication technology, project work, micro-teaching, field work, and self-reflection" (p. 61). At the end of the program, the graduates (who are now inservice teachers) are awarded certificates that qualify them to carry out educational research directed to promoting quality of learning, and to porbe creative ways to improve their teaching practice in schools based on education criteria for performance excellence in Bahrain.

By the end of the semester, the students who study a particular course at the University, including those who are enrolled in PGDE courses, are surveyed using a locally developed instrument to assess the teaching quality of that course. Results of student evaluations are regularly used to assess students' knowledge and skills gained within the course of study, and to provide the faculty with constructive feedback to improve teaching and learning (Al-Musawi, 2007). The subsequent evaluations of the quality of professional development programs and of the knowledge and skills of their graduates, undertaken by the Quality Assurance and Accreditation Center (Mohieldin, et al., 2010), an independent body that monitors the quality of teaching and learning by collecting documented evidence to capture the extent to which the Program's Intended Learning Outcomes (ILO) are compatible with a set of indicators of student performance, have not, however, produced satisfactory results in terms of the PD graduates' ability to transfer the knowledge about evaluation they had gained at the University into practical knowledge needed for effective teaching. Also, an examination of samples of students' works clearly demonstrated their lack of some important skills, such as the ability to build assessment tasks, to apply assessment methods in different contexts and to make sound, standard-based judgements based on objective criteria of achievement.

In the primary and secondary schools of the Kingdom of Bahrain, the use of assessment for student accountability focuses more on determining whether students have met various curriculum objectives, the criteria for a given

school curriculum or merit placement in a certain learning group within a class (Ministry of Education, 2012). The certification of students in Bahrain is largely a secondary school activity during the final three years of schooling and there are many significant consequences for individuals dependent on their performance on such assessments, including retention in a year or grade level, graduation, and tracking or streaming. Altogether, the said uses of assessment instantiate a conception wherein assessment is used as a means of making students accountable for learning. The prevailing conception of assessment, however, among the prospective teachers stems from the belief that learning objectives and test items determine what to teach (Al-Musawi, 2003, p. 38), and this conception is mainly associated with performance orientation and the use of various methods that would enable the teacher assess the student's work in a fair manner. Despite the fact that the teacher's orientation to assessment generally complies with the main requirements of the agenda of educational reforms in Bahrain, it does not, however, help in-service teachers to reveal the impediments and the strengths of student learning that would motivate then to find the optimal ways to improve student learning and the quality of instruction. This was a clear motive to investigate the possibility of using the conceptual change approach to change the teachers' conceptions on assessment and, subsequeently, their assessment practices.

As conceptions of teaching have been found to relate to teachers' approaches to teaching and also to students' approaches to learning, it is therefore logical to speculate that a change in conceptions of teaching of prospective and in-service teachers enrolled in the professional development programs at University of Bahrain would likely to be accompanied by a change in their evaluation practices and would eventually have an effect on the way their students approach studying in schools.

Based on this understanding, the research questions of the study are:

1- What is the impact of the staff development program for in-service teachers at the University of Bahrain on the conceptions of assessment of the participants?
2- What is the impact of the developed staff development program on the teaching and, consequently, assessment practices of in-service teachers at Bahrain University?
3- What is the effect of the developed staff development program on student ratings of the evaluation practices of in-service teachers at the University of Bahrain?

REVIEW OF LITERATURE

The importance of aligning teacher's assessment practices with their conceptions about assessment was investigated in a number of previous studies. Using cluster analyses, James and Pedder (2006) demonstrated that both the similarities and differences exist between the prospective teachers' values and practices in terms of their conceptions of assessment. Specifically, the teachers in their sample placed a high value for promoting learning autonomy in student assessment whereas they placed low value for the performance orientation, indicating that the teachers were aware of the importance of assessment for learning. In a similar study, Winterbottom et al. (2008) also found that the prospective teachers valued promoting learner autonomy more than they implemented in their teaching, whereas they applied performance orientation in the class more than they valued.

In another study, Vandeyar & Killen (2007) demonstrated that different views on assessment could lead to different assessment practices. Assessment will be utilized as an integral part of teaching and learning if educators view assessment as a useful means of gathering information on which important learning decisions will be based. Educators who see assessment as something that could be used to hold learners accountable for their own learning will favor formal, summative and high stakes assessment. Educators who view assessment as necessary but not important will favor summative and quasi-formative assessment practices that help in generating grades for reporting purposes. Educators who view assessment as largely irrelevant will probably avoid formative assessment in the classroom.

Based on a qualitative research design, Wang, Kao, & Lin (2010) examined the

Taiwanese prospective teachers' conceptions about assessment of science learning and the extent that these conceptions were coherent with their views of learning science. They found that the prospective teachers' conceptions of assessment can be identified in six categories: content knowledge; process of inquiry; attitude toward learning; measurement; performance; and informal assessment. Whereas their conceptions of learning can be identified using two categories: traditional tendency and constructivist tendency. Results showed that the prospective teachers reflected a traditional view of learning but held a more constructivist view about the methods of assessment, indicating a gap between their conceptions of assessment and learning.

Eren (2010) examined the consonance between prospective teachers' values and practices in terms of their conceptions and beliefs about assessment on a sample of 304 prospective teachers in the primary and secondary schools of Turkey. Results showed that the prospective teachers valued constructivist teaching and learning, making learning explicit, and promoting learning autonomy more than they practiced, whereas they practiced traditional teaching more than they valued.

In a recent study, Kitiashvili (2014) investigated the teachers' attitudes toward

assessment of students' learning and their assessment practices in Georgia's general educational institutions. The study included 106 schools, 928 surveyed teachers, and 292 observed classes, and the teachers, overall, had positive attitudes toward using assessment methods that require cognitive complexity from students. Although teachers' attitudes are related to their willingness toward using such assessment approaches in their classes to some extent, this study revealed the lack of correspondence between teachers' attitudes and their assessment practices.

The notion of changing teachers' attitudes to assessment appears also in Gareis and Grant's (2014) study, where an approach to developing the assessment literacy of teacher candidates in a public university was described. A definition of assessment literacy and a conceptual framework for the foundational knowledge and skills of assessment literacy were presented and explained within the context of a focused, one-credit course for pre-service general education teachers. Evidence of impact is provided, as are limitations and cautions. The article concludes with grounded insights into the need to develop the assessment literacy of teacher candidates.

With respect to the use of staff development programs to improve teachers' conceptions of assessment, Ho, Watkins, & Kelly (2001) explored the impact of the conceptual change staff development program on the conceptions about teaching held by twelve teachers in a University in Hong Kong. Results showed that the program brought about detectable conceptual change in two-thirds of the sample group. All the teachers who changed their conceptions on teaching received better ratings on their teaching practices from their students in the following academic year while none of those who did not change their conceptions showed similar gains in student rating scores. A positive impact on their students' studying approaches was also observed for half of the teachers who changed their conceptions on assessment.

In Ross, Ertmer, & Johnson (2001) study, thirteen school teachers participated in a technology integration professional development course taught in private schools in a Midwest city. The aim was to explore the effect of the staff development model, including peer modeling; peer collaboration; and reflection in an authentic learning context, on the beliefs and practices of the participants. The subsequent interviews with the teachers indicated that some of them revised their beliefs on the role of student but still expressed concern about assessment related to technology use.

Weurlander and Stenfors-Hayes (2008) investigated the effect of an innovative staff development course on Swedish medical teachers' thinking and practice. The analysis of the semi-structured interviews with 130 participants on the course one year after their participation showed that the course have contributed to a change in teachers' approaches to teaching and learning and their practice that ranged from using new technologies to basically transformed views about learning.

Iqbal, Azam, & Abiodullah (2009) widely discussed the assessment techniques practiced by teachers at higher education level, the way they implement assessment techniques and how effectively they integrate assessment in instruction for student learning. Thirty seven teachers of a metropolitan university in Pakistan took part in this study, and the results revealed that out of a vast variety of assessment techniques available to them, teachers use only very few of these. Authors concluded that higher education teachers need professional support for developing assessment strategies, building assessment literacy, and helping students to improve their learning.

Robinson, Reed, and Strauss (2011) sought to examine the impact of teachers' participation in an extended period of embedded professional development (PD) emphasizing teachers' use of assessment for learning (AFL) practices in extended problem based units of instruction within a comprehensive AFL framework. The comprehensive approach of using both course work and learning teams provided them with an in-depth understanding of these practices and the opportunity to apply that understanding in their daily work with their students. The assessment practices of the targeted eight elementary grade teachers changed dramatically as a result of the PD provided. Those changes were maintained even after external support was withdrawn, and teachers intended to continue the use of AFL strategies in the future.

The studies reviewed provide significant evidence and support for the impact of the staff development programs on teachers' conceptions on teaching and learning. This study, however, differs from the previous ones in three aspects: First, it uses the conceptual change approach to determine the effect of a staff development program on the change in conceptions on assessment; second, it is not restricted to a limited number of concepts and skills or to a specific aspect of assessment, such as table of specifications (Gareis and Grant, 2014); and, third, it is conducted on a sample of in-service teachers deeply involved in a professional development program in Bahrain.

METHOD

Design of the Evaluation Study

Based on the model for conceptual change programs (Ho, 2000), a short course of ten three-hour sessions scheduled in five consecutive weeks, with two sessions in a week, was designed for the purpose of the study. The non-credit program aimed to produce changes in the conceptions of assessment of the participants towards one that would be more conducive to quality teaching and student learning.

The program included the following elements: the self-awareness process; the confrontation process; exposure to better, alternative conceptions; and the commitment building and refreezing. The activities of the program in each stage are:

The Self-awareness Process: This stage was designed to engage participants in active reflection and critical analysis of their espoused conceptions and beliefs of assessment and their actual assessment practices. They were provided with a list of thought out questions that covered fundamental issues in assessment and were then asked to look at what they were actually doing in terms of assessment methods, test questions, etc. The Participants' written answers to the self-reflection questions were documented and were used in a later activity to analyze critically their own conceptions of assessment.

The Confrontation Process: When participants became aware of their own conceptions about assessment, they were confronted with the possible discrepancy that might exist between their espoused conceptions about assessment as recorded in their answers and their actual assessment practices, and the inconsistencies that might exist between their aspirations and their current beliefs. A series of organized activities was carried out to instill in participants an awareness of the urgent need to resolve these inconsistencies.

Exposure to Alternative Conceptions: Participants were provided with a conceptual framework that could help them to carry out a systematic analysis of their views about assessment and were exposed to better,

alternative conceptions on assessment that would challenge them to reconsider their ideas and beliefs. Presentations and group discussions were utilized to help them understand the meaning of the conceptions. This act facilitated their understanding of the factors that are more conducive to effective and deep learning, and those that are likely to induce surface learning for students.

Commitment Building and Freezing: Conceptions of assessment alternative to the ones participants were espousing were created. Teachers were engaged in planning changes for their teaching so as to encourage them build up a commitment to change.

Measuring the impact on assessment conceptions and practices:

The impact of the staff development Program of study on the participants' assessment conceptions was defined by identifying and comparing their conceptions of assessment before and after the Program. In order to solicit conceptions of assessment espoused by the participants, three, semi-structured interviews were employed: The Pre-Program Interview recorded their conceptions on assessment before they attended the Program; The Immediate Post-Program Interview identified the urgent impact of the Program; and, lastly, the Delayed Post-Program Interview conducted one year after the Program allowed more lasting changes to be clearly differentiated among the teachers.

To detect changes in assessment practices and student learning which might be a consequence of participation in the Program, a longitudinal study that required the participants to be followed for a pre-Program year and a post-Program year was set up. While the academic year 2013/2014 served as the 'control year', the following year 2014/2015 was taken as the 'test year' when the teachers had completed the Program. Each participant selected a school subject which he or she taught in both control and test years for use in this evaluation study. The use of the same subject allowed the nature of conceptions on assessment (Brown, 2008) to be controlled for.

The Bahraini version of the Course Experience Questionnaire (CEQ) (Ramsden, 1991), developed by the author (Al-Musawi, 2013) was used to measure the effect of the Program on the teaching practices of teachers as perceived by their students. By comparing the CEQ scale scores for the test-year and that for the control-year of the practices, a deduction can then be made to any possible influence of the Program.

The consequential effect of the Program on student learning was concluded by comparing the impact of the participant's teaching on the studying approaches of their students in the test and the control years using the Revised Approaches to Studying Inventory (RASI) (Entwistle & Tait, 1995). In each year then, the RASI was administered to students at the start and at the end of the Program.

Participants

The sample consisted of twelve (4 males and 8 females) in-service teachers enrolled

in their first year of study in the PGDE Program at the University of Bahrain, and working at the same time in secondary schools of Bahrain, majoring in teaching science (n=4), mathematics (n=5), and

English language (n=3). Participants' age ranged from 22 to 26 years (Mean = 23.4 years).

Instruments:

1- The Interview:

Before the interview, the instructor contacted the school principals and obtained permission to interview the teachers. During the interview, the researcher acted as a facilitator, posing questions and adding prompting follow-up questions, encouraging the teachers to elaborate on certain points and offer additional comments. The interview went smoothly, and the teachers seemed more relaxed when they heard the introduction and learned about the interview questions at the start of the interview. While most of the time the views seemed to be shared by the majority of the interviewees, occasionally some clearly different opinions were exchanged at the interview. Immediately after the interview, the researcher summarized the group's responses to the questions, and sent the summary to the teachers to verify if the summary truthfully reflected what was talked about during the interview.

The interview questions (see Appendix 1) were developed with reference to relevant literature about conceptions of assessment (Brown, 2004; Eren, 2010). The questions aimed to solicit in-service teachers' espoused conceptions on assessment and their practices. Interviews were conducted, audiotaped and transcribed verbatim by the author. Before the interview, the teachers were reminded that all questions should be answered with reference to the subject matter they teach in their schools.

To conduct the interview, the author undertook the following steps widely used in the field of qualitative research (Carey & Asbury, 2012; Farnsworth & Boon, 2010):

1. Preparation: The author was fully prepared prior to the interview. He visited the teachers who should be interviewed ahead of time, read the resources about them and their schools to gain information, and made notes when researching the topic in the proper educational databases to form the questions to be asked during the interview.
2. Forming Questions: Before the interview begins, the author created questions that targeted the teacher's educational background and work experience, the training workshops and sessions related to assessment issues that he or she attended at work.
3. Showing Courtesy: Since the interviewee is giving time out of their schedule to help with this research, the participating teacher deserves to be shown respect, a fact that prompted the author to arrive on time for the interview, politely greeting the interviewee and introducing himself to each of the teachers by name and position.
4. Engagement and Explanation: The process for the interview was explained and informed consent was received from all teachers to ensure that he or she understands the purpose of the interview and is comfortable with discussing his or her knowledge about assessment and the methods he or she uses to assess students. Before the interview begins, the author asked the interviewee if the session can be recorded and gained the acceptance from each of the interviewees, because a tape recording of the session is very helpful to listen to while writing out the results of the interview.
5. The interview process: The author started the interview by allowing the teacher to describe how he or she perceives assessment in terms of definition, purposes, methods used in the class, their strengths and weaknesses and to give any additional information that he or she feels relevant to the topic, without being interrupted. Teachers were encouraged to explain their perspectives on the topic and were given the time needed to fully share their teaching and learning experience, while the author carefully documented the information as appropriate. A positive relationship was maintained between the author and the teachers who were really comfortable with the interview questions and answered them with an acceptable level of clarity.
6. Follow-Up Questions: While listening to the teacher's answers to the interview questions, the author asked additional open-ended questions that allowed the teacher to describe in great detail on the focus of his or her assessment in the classroom, what he or she usually seeks to assess in exams, quizzes, and projects, and what he or she does to achieve congruence between the test questions and lesson objectives. The interviewees were also urged to provide examples and incidents of assessments to corroborate their ideas and conventions. Specific, closed-ended questions that can be answered by 'Yes' or 'No' were also used last to complete any missing information. For any topic discussed during the interview, the author clarified the terms and definitions used by the interviewee to ensure accuracy of meaning and interpretation.
7. Ending the Interview: The author ended the interview on good terms, asking the participating teachers whether they have any other information to add or share with him, then he thanked them for their time, information and cooperation with him.
8. Transcribing the Interview: The interviews were transcribed immediately while the

facts remained fresh in memory to keep the information as accurate as possible. The transcripts of the three interviews were analyzed for indications of teachers' conceptions of assessment in the different stages of the staff development program and were compared to reveal any changes that might have occurred. To allow a methodical analysis, Brown's dimensions of assessment were chosen as a reference framework; that is, assessment improves teaching and learning (Improvement); assessment makes schools and teachers accountable (School Accountability); assessment makes students accountable (Student Accountability), and assessment practices that are irrelevant (Irrelevance); and that respond to public examinations (Examination) (Brown, et al., 2009). A preliminary analysis of the interview transcripts revealed that the conceptions of assessment presented in the interview data set of this study largely fell within the selected theoretical framework above. Table 1 displays the final scheme of the five dimensions and their representative attitude codes (A and B) used in the indexing of interview transcripts.

The Thematic Structural Analysis (TSA) was used to index the transcripts while being carefully read to determine whether the respondent was expressing opinions related to any of the five dimensions, and toward which attitude pole the opinion was inclined. The TSA method (Alexander, Muir, & Chant, 1992) stems from the central idea in content analysis that many words in a given set of

Table 1
Dimensions and Attitude Code used in the Analysis of the Interview Data (Adapted from Brown, et al., 2009)

Dimensions	Attitude Codes
1- Improvement	**A** Describing student learning and performance
	B Diagnosing and informing student learning
2-School Accountability	**A** Demonstrating quality instruction/learning
	B Improving quality of instruction/learning
3- Student Accountability	**A** Assigning grades against criteria of success
	B Determining levels against quality standards
4- Irrelevance	**A** Being inaccurate and unfair to students
	B Lacking vital information for students
5- Examinations	**A** Measuring memorization of content
	B Measuring higher-order thinking skills

texts can be classified into many fewer content categories (Weber, 1985). Using this technique helped search for the central ideas implicitly included in the teachers' answers to explore their beliefs.

After the analysis of interview data, the twelve participants were placed into 'change groups' according to their degree of conceptual change, or in equivalent terms, degree of changes in their conceptions of assessment, as indexed by TCA, and degree of changes predicted for their students, and CEQ and RASI results were then analyzed for each of the groups. The teachers were placed into the following groups: The 'Yes', the 'Unsure', and the 'No' change groups, where the positive changes in the CEQ and RASI results between the two year groups of their students were likely to happen; not sure to be likely to happen; and unlikely to happen, respectively. Five faculty members helped classify the teachers into the 'change groups'.

2- The Course Experience Questionnaire (CEQ):

Students' perception of the teaching provided by a teacher was measured by the five scale scores in the CEQ, which cover the following

aspects: Good Teaching, Clear Goals and Standards, Appropriate Workload, Appropriate Assessment and Emphasis on Independence. The CEQ scales had adequate reliability and factor structure for research purposes with this sample (Al-Musawi, 2013).

The 25 items of the instrument used in this study were selected from the original long form of the CEQ (CEQ30) (Wilson et al., 1997) on the basis of the content of the Higher Diploma in Education programme. 18 of the original items of the CEQ30 were retained in the same form, and the rest were partially rephrased or completely rewritten so as to reflect the specific features of the local teaching and learning environment surrounding the PD programs taught at Bahrain University.

In the Wilson, et al. 1997, question 37 asked about the students' overall satisfaction with the quality of the course. In the present study (CEQ25) this question was numbered 25 and was aimed to assess the students' overall satisfaction with the teaching quality of the course. Each item contained a statement requiring a response using a 5-point Likert scale, ranging from a score of 1(strongly disagree) to 5 (strongly agree) to reflect students' views of the program as a whole (not with reference to a certain teacher). Consequently, the total score ranged from 25 to 125.

A comparison of the scale scores between the pre-Program year and the post-Program year gave an indication of how differently the students perceived the classroom teaching of their own teachers. One-way ANOVA was carried out for each change group of in-service tecahers, with student cohort (control-year vs test-year cohort) as a between-subject variable.

3- The Revised Approaches to Studying Inventory (RASI):

The approaches to studying of the students taught by the teachers were represented by the six scale scores of the RASI: deep approach; surface approach; strategic approach; lack of direction; academic self-confidence and meta-cognitive awareness of studying (Entwistle & Tait, 1995). Differences in the RASI scale scores between the pre- and post-test within the same year was taken as an indication of the impact of the teaching on students' learning approaches. One-way Analysis of Variance was used to analyze the changes in the six RASI scale scores in the beginning and the end of the Program.

The Arabic version of the Revised Approaches to Studying Inventory (RASI) was prepared by the author. Utilizing the methodology suggested for development of translated tests (Brislin, 1986), the original, English version of the RASI was translated into Arabic, and the cross-language equivalence of the two forms of the RASI was achieved successfully. The measure was then administered to a sample of 50 in-service teachers involved in the PGDE program, and its factor structure was examined using exploratory factor analysis. Results of the validation study of the Arabic version of RASI confirmed its suitability for the sample in this study.

Data Analysis

In this study, both qualitative and quantitative data were gathered and analyzed; therefore, the results of a mixed-method research are presented. The qualitative data was collected via semi structured in-depth interviews. To analyze and interpret the qualitative data, the Thematic Structural Analysis (TSA) was used. The data collected via the interviews were first transcribed, translated into English and categorized. The quantitative data collected from both the Course Experience Questionnaire (CES) and the Revised Approaches to Studying Inventory (RASI) were statistically analyzed using the Statistical Package for the Social Sciences (SPSS), version 20.0. The descriptive analysis was performed to find means and standard deviation and the ANOVA was used to compare results with the criteria.

RESULTS

The effect of the Program on conceptions of assessment. The twelve program participants were classified into three 'change groups'. Highly consistent results were obtained from the panelists as the Hoyt estimate of reliability averaged over the five raters was 0.91, an indication of high inter-rater reliability.

As far as two teachers from the 'Yes' group were concerned, the majority of the

raters were convinced that the changes in these teachers' conceptions of assessment were significant enough to have positive effect in their teaching and assessment practices and to bring about corresponding changes in their students' learning. Another two teachers were put in the 'Unsure' group, indicating that the panelists acknowledged noticeable changes in their conceptions of assessment. For the other eight teachers, the raters found only minor or no changes in their conceptions about assessment. The results therefore suggested that a success rate of four teachers was achieved for conceptual change, though the degree of change varied among them.

The effect of the Program on assessment practices

The means and standard deviations of the CEQ scale scores obtained in the pre-Program (control) year and the post-Program (test) year for the different groups of teachers are shown in Table 2.

Table 2
Means and Standard Deviations of CEQ Scales by Teacher Group

CEQ Dimension		Teacher Group 'Yes'	Teacher Group 'Unsure'	Teacher Group 'No'
Good Teaching	Control year	1.93 (.83)	3.25 (.74)	2.25 (.59)
	Test year	3.90 (.90)	3.39 (.61)	2.28 (.69)
	Difference	1.97	.14	.03
Clear Goals	Control year	1.92 (.83)	3.26 (.58)	3.28 (.66)
	Test year	4.08 (.92)	3.15 (.42)	3.26 (.51)
	Difference	2.16	-0.11	-.02
Appropriate Workload	Control year	1.95 (.78)	3.74 (.81)	3.47 (.52)
	Test year	3.95 (.96)	3.61 (.76)	3.43 (.64)
	Difference	2.00	-.13	-.04
Appropriate Assessment	Control year	1.78 (.80)	3.51 (.53)	2.82 (.60)
	Test year	4.25 (.84)	3.25 (.45)	2.86 (.58)
	Difference	2.47	-.25	.04
Emphasis on Independence	Control year	1.83 (.78)	3.19 (.67)	3.94 (.57)
	Test year	4.25 (.81)	3.28 (.72)	3.99 (.63)
	Difference	2.42	.13	.05

Note. Difference = test year mean – control year mean; values in parenthesis are standard deviations.

The comparison of the CEQ scores of the 'Yes' change group before and after the Program confirmed that the teachers in this group achieved an improvement in students' ratings on their teaching after attending the Program. Significant differences were found on all CEQ scales in the univariate analysis: Good Teaching, $F(5.18) = 11.29$, $p<0.01$, Clear Goals, $F(5.18)= 7.80$, $p<0.01$, Appropriate Workload, $F(5.18)= 28.59$, $p<0.01$, Appropriate Assessment, $F(5.18)= 8.52$, $p<0.01$, and Emphasis on Independence, $F(5.18)= 9.18$, $p<0.01$. The 'Unsure' change group showed a moderate yet insignificant improvement in students' perceptions of their teaching in the post-program year, while the 'No' group demonstrated no improvement in student ratings.

The effect of the Program on students' approaches to studying

The means and standard deviations of

the RASI scale scores are displayed in Table 3. The 'Yes' change group improved in the Post-Program year in terms of the influence on students' studying approaches as teachers were able to facilitate students towards a deep approach to studying.

Table 3
Means and Standard Deviations of RASI Scales by Teacher Group

RASI Dimension		Teacher Group 'Yes'	'Unsure'	'No'
Deep Approach	Control year	1.78 (.86)	3.12 (.75)	2.76 (.54)
	Test year	3.18 (.90)	3.24 (.82)	2.74 (.46)
	Difference	1.40	.12	-.02
Surface Approach	Control year	1.90 (.90)	3.21 (.49)	3.88 (.59)
	Test year	3.10 (.84)	3.39 (.81)	3.93 (.71)
	Difference	1.20	.18	.05
Strategic Approach	Control year	1.95 (.88)	3.11 (.87)	3.94 (.67)
	Test year	3.10 (.84)	2.99 (.65)	3.87 (.89)
	Difference	1.15	-.12	- .07
Lack of Direction	Control year	1.70 (.97)	3.71 (.63)	2.27 (.57)
	Test year	2.90 (1.08)	3.46 (.62)	2.28 (.81)
	Difference	1.20	-.25	.01
Academic Self-confidence	Control year	1.83 (1.06)	3.29 (.69)	3.17 (.43)
	Test year	2.78 (1.07)	3.43 (.77)	3.27 (.55)
	Difference	.95	.14	.10
Meta-cognitive Awareness	Control year	1.97 (.86)	3.86 (.61)	3.68 (.34)
	Test year	2.93 (.85)	3.67 (.66)	3.64 (.62)
	Difference	.96	-.19	-.04

Note. Difference = test year mean – control year mean; values in parenthesis are standard deviations.

Univariate analysis indicated significant differences for all the six RASI scales: Deep approach, $F(5.18)= 8.70$, $p<0.01$; surface approach, $F(5.18)= 5.41$, $p<0.01$; strategic approach, $F(5.18)= 4.54$, $p<0.01$; Lack of direction, $F(5.18)= 3.71$, $p<0.01$, Academic self-confidence, $F(5.18)= 8.45$, $p<0.01$, and Meta-cognitive awareness of studying, $F(5.18)= 6.84$, $p<0.01$. The 'Unsure' change group showed a significant improvement in Academic Self-confidence dimension, $F(4.79)= 7.29$, $p<0.01$, while no significant effect was obtained for the 'No' change group.

DISCUSSION

The results of the study corroborate the findings of the previous studies (Ho, et al., 2001; Watkins & Kelly, 2001; Robinson, et al., 2015) which showed a positive effect of a staff development program on conceptions of assessment, teaching practices, and students' approaches to studying. Though the

scope of the current study did not allow for examination of the impact of the Program on individual in-service teachers, it can be assumed that the results for members of each change group would support those obtained for all teachers involved in that group.

The findings from this small sample of pre-service teachers are consistent with the findings from a larger study of pre-service teachers which was conducted in South Africa (Vandeyar & Killen, 2007). This study found that teachers' perceptions of assessment could lead to different assessment practices. The results are also in line with the findings of Dayal and Lingam's (2015) study that the majority of teachers initially provided a narrow understanding of assessment but many teachers from this group, however,

were able to identify formative functions of assessment, regardless of the number of years of teaching experience, years of training and teacher's gender.

Overall, teachers' conceptions about assessment have not changed dramatically as a result of the professional development program provided. Although in-service teachers reported an increased level of understanding of assessment, the relationship between measurement, assessment, and evaluation, and the aims of evaluation for the subjects they are teaching in schools, they were still unaware of what they are actually trying to assess in their exams, tests, quizzes, projects, and how these assessment tools can be used effectively to promote student learning in the classroom. Following the PD program, most participants had problems in moving from the traditionalist to the constructivist view of assessment as an efficient tool to diagnose and inform student learning and to improve the quality of instruction and learning and of exams as a mean for measuring the higher-order levels of thinking skills. Rank ordered from most to least, the greatest differences in in-service teachers' conceptions on assessment were in how to determine levels of student achievement against quality standards, diagnose and inform student learning, tap higher-order skills, and improve the quality of classroom instruction in line with the guiding directions of quality assurance agencies in Bahrain.

When the researcher first asked the teachers who participated in the in-depth interviews what the word "assessment" conjures up in their mind, most of the participants described assessment as an effort by instructors to check the status of student learning. They mentioned homework, oral queries, classroom activities, quizzes, tests, and projects. During this discussion, the more experienced members of the participants offered a different view. To these teachers assessment as a process of gathering information about student progress in the subject differs from the process of evaluation that entails making certain judjments and decisions about the collected information, a task that often goes beyond the individual teahcer's competence. At that moment of the interview, the researcher asked if the participants felt the definition of assessment had shifted in recent years. Several teachers agreed that the term assessment now does not just mean grading students against the school criteria of achievement, but also designing the classroom assignments in a way that enables the student to practice certain skills in the classroom, making it more of a learning activity instead of an assessment activity. As a result, the classroom assignment functioned as an opportunity for students to learn and to move forward toward the realisation of lesson objectives. This, in turn, enabled the instructor to check how well students could apply their learning during this practice. As to the relationship between assessment and teaching, some teachers said that the two processes were often intertwined. They reasoned that assessment would support instruction only if the two occur almost simultaneously. One teacher explained that if a formative assessment showed that her students had not understood a concept, she would use a new example to explain that concept. This adjustment in instruction should follow an assessment activity immediately to make the assessment worthwhile (Yao, 2015).

As to the impact of the professional development program on the teaching and, consequently, assessment practices of in-service teachers, the extended and comprehensive approach of allowing the participating teachers to be in direct, face-to-face with the possible discrepancy that might exist between their espoused conceptions about assessment as recorded in their answers to interview questions and their actual assessment practices, had provided them with an indepth understanding of these practices and the opportunity to apply that kind of understanding in their daily work with their students. It is not surprising that in-depth training coupled with the exposure to alternative conceptions over an extended period of time would lead to such changes, but insights gained from this study provided additional dimensions to consider in providing effective professional development in the use of practices associated with the new understanding of assessment for learning, not of learning.

Furthermore, teachers in the same grade level need to have an in-depth and similar

background in their assessment practices in order to work productively together. When they don't, teachers cannot engage in the kind of thoughtful dialogue needed to move them forward in crafting ideas into concrete classroom practice. While much is now known about the effective use of formative assessment, there is a need to better understand how other contextual factors such as students' approaches to studying and the inservice teachers' conceptions of teachers working together interact with these initiatives to either promote or subvert the process of assessment.

While inquiring about the teachers' assessment practices, the researcher asked them what parts of the classroom process assessment activities were and what parts instructional activities were. The participants responded that assessment took place all along in this process. It happened when the students performed the classroom assignments and also when the instructor demonstrated a new process or queried the students as a whole group, or individually. It also happened when the students did the various classroom activities guided by the teacher, and the independent homework assignments. The interviewees suggested that they assessed students in multiple ways, including asking students individually or as a whole group, and walking near the students to check their progress as a group or as they work individually. Although the most experienced teacher defined assessment as traditional testing, he did mention that he had his own ways of checking student understanding other than regular tests, that is by watching the eye movements of stdents who were having difficulty with the lesson but not frequently enlisting the teacher's help and support.

The in-service teachers considered giving students assignments that were project- based as a real learning activity, because these assessments provided an opportunity for students to apply their learning and solve real world problems. They believed that students also enjoyed completing such projects. This view seemed to be shared by several other teachers in the group. As to group activities that students were asked to participate in during the lesson, the teacher who mentioned this type of assessment said that the interactions allowed the students to explain, to ponder, to assess, and to learn from each other in the class. In such assessment activities, his students were often actively engaged. Other teachers, however, proclaimed that the test is the most important tool of assessment because that was when students could reveal to them what they had learned, not to say that the test results allows the teacher to reflect on his or her work with students, give the students the necessary feedback on tha learned material, and find various ways to motivate them to improve their own learning.

When asked what aspect of assessment presents the biggest challenge for them, the teachers believed it was not easy for them to write or choose effective test questions, especially questions that were used for standardised assessments for schools located in one region, because these questions require substantial time and effort on the part of the teacher, and because of difficulty of matching a test question with the lesson objectives on the one hand, and the subject matter knowledge and skills as predescribed in the formal educational documents, on the other hand. In general, these findings corroborate the results of very recent studies that inservice teachers' lack the testing skills necessary for effective learning (Abed & Awwad, 2016, p. 75).

Whith respect to the effect of the staff development program on the student ratings of the evaluation practices of in-service teachers, it is worth mentioning that the majority of the teachers in the focus group expressed their satisfaction when their students got the opportunity to rate their use of tests, homework assignment, and projects as effective tools of assessing student learning. While the evaluation of a particular course or lecturer, as Desith (2007) points out, "does not measure how the learning environment as a whole affects the students", it does, however, "provide more specific, focused, and potentially useful information about the strengths and weaknesses of a particular lecture" (p. 186), and, hence, about the context of learning.

This result could be attributed to two main factors, first, the positive change in teaching and assessment toward the constructive views

of learning and assessment on the part of service teachers; and second, the positive change of students' approaches toward the deep and strategic approaches to learning . In the context of the concept of student's approaches to learning being the intentions and motives a student has when approaching a learning task, as well as the corresponding strategies by which these intentions and motives are accomplished, students of the in-service teachers who participated in the current study started to gradually embrace a "deep approach" to learning that refers to "an intention to understand, with use of evidence and relating of ideas as the predominant strategies, and interest in ideas as the predominant motive", and also a "strategic approach" that refers to a student's intention "to succeed and the motive to achieve the best grades possible, by organization of time and the learning environment", as opposed to the "surface approach" to learning, where the student concentrates only on "reproducing the learning material by means of rote learning to avoid failure" (p. 188). In this sense, it is important to note here that the student's approaches to learning reflect his or her perception of the learning environment, rather than the environment in an objective sense, which influenced learning. Within the context, This fact clearly demonstrates the effectiveness of the professional development program in bringing about vital detectable changes in the students' perception of learning and assessment, which ultimately changed their approaches to learning in an invitational positive manner.

While it is difficult to generalize the results of the study given the small number of participants, the impact of the conceptual change approach on teachers' conceptions of assessment and on their assessment practices is obviously evident: four of the twelve teachers involved in the Program managed to change their conceptions on assessment and their corresponding students consequently received higher grades in the following year, while none of the students of those teachers who did not change their conceptions of assessment showed similar gains in achievement. Such an outcome implies that the change in teachers' conceptions about teaching in general, and assessment as a major component of teaching in particular, is likely to happen when the teacher developer maintains some consistent focus and message while working in a close relationship with a small number of teachers (Al-Musawi, 2002), as it was observed in this study.

The fact that the majority of teachers did not change their views about assessment might be attributed to the short duration of the Program because if teachers are to acquire a non-traditional orientation, curriculum messages must be reinforced and supported in all the facets of a long-term and effective program of change that encompasses all aspects of teaching and learning (Graber, 1996). This is the reason why the anticipated change did not happen although the teaching objectives of the professional development programs at the University strive to foster students' ability to "select the appropriate assessment methods that really influence student learning and to use them effectively in the classroom" (University of Bahrain, 2010, p. 34).

The likelihood of the conceptual change to take place among in-service teachers largely depends, then, on how they understand the theoretical foundations of assessment and how they effectively apply them to promote quality learning in their own schools. As the current methods of instruction and assessment in tertiary education are mostly traditional and inadequate to the student's particular needs and interests, the teacher can hardly find sufficient time to inquire and to reflect upon what is being taught (Al-Musawi, 2003). This can explain why some participants expressed mixed orientations on assessment in the end of the staff development program, a fact which necessities aligning teachers' assessment practices with their beliefs and values about teaching and learning (Borgioli1, Ociepka & Coker, 2015), and building of a strong evaluative culture that "seeks out empirical information on its performance in order to use that information to learn how to better manage its programs and services, and thereby improves its performance" (Mayne, 2009. p. 6).

Based on this understanding, the main learning occurs during the process of the evaluation, and hence if managers and staff are involved in the process of measuring

and analyzing results information, they are more likely to make use of the information gathered and to build interest in an evaluation culture. This, however, may happen only if the evaluation is perceived as a participatory process that "captures the spirit of change, which is the very essence of human development" (Crishna, 2006, p. 218), rather than just a tool for achieving accountability in institutions of higher education.

CONCLUSIONS AND IMPLICATIONS

Results of the study shows that the teachers' assessment conceptions and school practices have undergone limited but clearly observable change as a result of the professional development program provided. The extended course and mixed method approach of using both interviews and evaluation instruments had provided teachers with an indepth understanding of assessmnet practices and the opportunity to apply that understanding in their daily work with their students. It is not surprising that in-depth training coupled with the use of learning experience over an extended period of time would lead to such changes but insights gained from this study provided additional dimensions to consider in providing effective PD in the use of assessment for learning. The biggest chalange for teachers is to translate the body of knowledge about assessment into meaningful classroom practice in the absence of the external control and pressure. Obviously, most teachers do lack the apportunity to implement the learned assessment methods in the daily classroom practice because the teachers' busy scheduals, limited time for planning, and need to cover the curriculum subvert such efforts, not to add that teachers in the same grade level need to have an in-depth and similar background in teaching in order to work together to practice formative assessment knowledge, skills and activities in the context of classroom environment.

As Greenstein (2010) pointed out, teachers and teacher candidates need support in developing those formative assessment skills, through professional development activities and teacher education programs. In particular, teacher education programs play an important role in ensuring that prospective teachers have a coherent view of classroom assessment, and a full understanding of the implications of assessment for learning. Although the interview questions posed at the focus group were focused on a range of topics related to teacher perceptions of classroom assessment, future research could focus on additional areas of teacher perceptions regarding assessment. Moreover, we believe that those responsible for in-service teachers' training should encourage teachers to use assessment results to help develop students, and the school administrators should specify that assessment results be used to improve learning and they must continually monitor teachers to make sure that it is carried out properly. It is also essential, in our opinion, that teacher development researchers convince school administrators of the need to provide adequate support, and that schools must bear responsibility and learning together with teachers, and must not let them carry out assessments on their own without considering the ultimate goals of teaching.

REFERENCES

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.

Airasian, P. W. (1997). Classroom assessment. New York: McGraw-Hill.

Alexander, D., Muir, D., & Chant, D. (1992). Interrogating stories: How teachers think they learned to teach. Teaching and Teacher Education, 8(1), 59-68.

Al-Musawi, N. (2002). The impact of student teaching programs on student beliefs' about teaching and learning. Journal of the International Society for Teacher Education (JISTE), 6(1), 10-19.

Al-Musawi, N. (2003). The effect of student teaching programs in the College of Education at the University of Bahrain on students' beliefs about teaching and learning processes. Kingdom of Bahrain: University of Bahrain Press.

Al-Musawi, N. (2007). Developing an instrument to measure effective university teaching. The Educational Journal (Kuwait University), 21(83), 61-98.

Al-Musawi, N. (2013) . Validation and use of the Course Experience Questionnaire to assess the professional development programmes at Bahrain University. The International Journal of Educational and Psychological Assessment, 13(1), 74-91.

Argyris, C. & SchÖn, D. A. (1974). Theory in practice: Incresing professional effectiveness. San Francisco: Jossey-Bass.

Bahrain Teachers College (BTC) (2009). Prospectus. BTC: University of Bahrain.

Biggs, J. (1989). Approcaches to the enhancement of tertiary teaching. Higher Education Research and Development, 8(1), 7-25.

Black, P. & Wiliam, D. (1998). Assessment and classroom learning. Educational

Assessment: Principles, Policy and Practice, 5(1), 7-74.

Borgioli1, G. M., Ociepka, A., & Coker, K. (2015). A playbill: Rethinking assessment in teacher education. Journal of the Scholarship of Teaching and Learning, 15(3), 68-84.

Bowden, J. A. (1989). Curriculum development for conceptual change learning: A phenomenographic pedagogy. Paper presented at the Sixth Annual Conference of the Hong Kong Educational Research Association, Hong Kong.

Brislin, R. W. (1986). The wording and translation of research instruments. In W. J. Lonner & J. W. Berry (Eds.), Field methods in cross-cultural research (Vol. 8.,

pp. 47-83). Beverly Hills, CA: Sage.

Brown, G. T. L. (2004). Teachers' conceptions of assessment: Implications for policy and professional development. Assessment in Education, 11(3), 301-318.

Brown, G. T. L. (2008). Conceptions of assessment: Understanding what assessment means to teachers and students. New York, NY: Nova Science Publishers.

Brown, G. T. L., Kennedy, K. J., Fok, P. K., Chan, J. K. S., & Yu, W. M. (2009). Assessment for student improvement: Understanding Hong Kong teachers' conceptions and practices of assessment. Assessment in Education, 16(3), 347-363.

Butterfield, S., Williams, A. & Marr, A. (1999). Talking about assessment: mentor-student

dialogues about pupil assessment in initial teacher training. Assessment in Education,

6(2), 225-246.

Carey, M. A., & Asbury, J. (2012). Focus group research. Walnut Creek, CA: Left Coast Press.

Carnell, E. (2007). Conceptions of effective teaching in higher education: Extending the boundaries. Teaching in Higher Education, 12(1), 25-40.

Charoenchai, C., Phuseeorn , S., & Phengsawat, W. (2015). Teachers' development model to authentic assessment by empowerment evaluation approach. Educational Research and Reviews, 10(17), 2524-2530.

Chen, H. T. (2005). Practical program evaluation: Assessing and improving planning, implementation and effectiveness. Thousand Oaks, CA: Sage.

Chen, H. T. (2006). A theory-driven evaluation perspective on mixed-method research. Research in the Schools, 13(1), 75-83.

Cirit, N. C. (2015). Assessing ELT pre-service teachers via Web 2.0 tools: Perceptions toward traditional, online and alternative assessment. The Turkish

Online Journal of Educational Technology (TOJET), 14(3), 9-19.

Crishna, B. (2006). Participatory evaluation (1) – Sharing lessons from fieldwork in Asia. Child: Care, Health and Development, 33(3), 217-223.

Crooks, T. (1988). The impact of classroom evaluation practices on students. Review of Educational Research, 58(4), 438-481.

Dayal, H. C., & Lingam, G. I. (2015). Fijian teachers' conceptions of assessment. Australian Journal of Teacher Education, 40(8), 42-58.

Desith, A. (2007). Students' evaluation of teaching, approaches to learning and academic achievement. Scandinavian Journal of Educational Research, 51(2), 185-204.

Entwistle, N. J., & Tait, H. (1995). The Revised Approaches to Studying Inventory. Edinburgh: Centre for Research on Learning and Instruction, University of Edinburgh.

Eren, A. (2010). Consonance and dissonance between Turkish prospective teachers' valuaes and practices: conceptions about teaching, learning, and assessment. Australian Journal of Teacher Education, 35(3), 27-48.

Farnsworth, J., & Boon, B. (2010). Analyzing group dynamics within the focus group. Qualitative Research, 10(5), 605-624

Gareis, C. R., & Grant, L. W. (2014). Assessment literacy for teacher candidates: A focused approach. Teacher Educator's Journal, 2015, 4-21.

Gibbs, G. (1995). Changing teachers' conceptions of teaching and learning through action research. In A. Brew (Ed.), Directions in Staff Development (pp. 21-35). Buckingham: Society for Research into Higher Education and Open University Press.

Gordon, M. (2008). Between constructivism and connectedness. Journal of Teacher Education, 59, 322-331.

Graber, K. C. (1996). Influencing teacher beliefs: The design of a 'high impact" teacher education program. Teaching and Teacher Education, 12(5), 451-466.

Greenstein, L. (2010). What teachers really need to know about formative assessment. Alexandria, VA: ASCD.

Guthrie, J. T. (2002). Preparing students for high-stakes test taking in reading, in:

A. E. Farstrup & S. J. Samuels (Eds.), What research has to say about reading instruction, Newark, DE, International Reading Association, pp. 370-391.

Hargreaves, E. (2007). The validity of collaborative assessment for learning. Assessment in Education: Principles, Policy, & Practice, 14(2), 185-199.

Hershberg, T. (2002). Comment, in: D. Ravitch (Ed.) Brookings Papers on

Education Policy: 2002. Washington DC, Brookings Institution Press, pp. 324-333.

Higher Education Policy Institute for the Education Reform Board (2010, March). Bahrain's Higher Education Agenda: A higher education improvement strategy. Kingdom of Bahrain: Ministry of Education Documents.

Ho, A. S. P. (1998). A conceptual change staff development programme: Effect as perceived by participants. International Journal for Academic Development, 3, 24-38.

Ho, A. S. P. (2000). A conceptual change approcah to staff development: A model for programme design. International Journal for Academic Development, 5(1), 30-41.

Ho, A. S. P., Watkins, D., & Kelly, M. (2001). The conceptual change approach to improving teaching and learning: An evaluation of a Hong Kong staff development programme. Higher Education, 42, 143-169.

Iqbal, H. M., Azam, S., & Abiodullah, M. (2009). Using assessment for improving students learning: an analysis of university teachers practices. Bulletin of Education and Research, 31(1), 47-59.

James, M., & Pedder, D. (2006). Beyond method: Assessment and learning practices and values. Curriculum Journal, 17(2), 109-138.

Kitiashvili, A. (2014). Teachers' attitudes towards assessment of student learning and teacher assessment practices in general educational institutions: The case of Georgia. Improving Schools, 17(2), 163-175.

Linn, R. L. (2000). Assessments and accountability. Educational Researcher,

29(2), 4-16.

Lorente-Catalàn, E., & Kirk, D. (2015). Student teachers' understanding and application of assessment for learning during a physical education teacher education course. European Physical Education Review, 1(1), 1-17.

Lorna, E. (2003). Assessment as learning: Using classroom assessment to maximize student learning. Thousand Oaks, CA: Crown Press.

Madina, A., Baker, E. L., Chow, K. A., Delacruz, G. C., & Griffin, N. C. (2015). Assessment of teachers from a social psychological perspective. Review of Research in Education, 39(1), 54-86.

March, H. W. (1987). Students' evaluations of university teaching: research findings, methodological issues, and directions for future research. International Journal of Educational Research, 11, 253-388.

Mayne, J. (2009). Building an evaluative culture: the key to effetive evaluation and results management. The Canadian Journal of Porgram Evaluation, 24(2), 1-30.

McMillan, J. H. (2007). Classroom assessment: Principles and practice for effective standard-based assessment (4th ed.). Boston, MA: Allyn and Bacon.

Ministry of Education (2012). A decade of development (2003-2012): Educational development documents. Kingdom of Bahrain, Ministry of Education, Directorate of Public Relationships and Media.

Mohieldin, T.,, Al-Ammal, H., & Al-Burshaid, M. (2010). IDEAS: A university-wide outcome-based assessment process. The Quality Assurance and Accreditation Center, University of Bahrain: University of Bahrain Printing Press.

Posner, G. J., Strike, K. A., & Hewson, P. W. (1982). Accomodation of a scientific conception: Toward a theory of conceptual change. Science Education, 66, 211-227.

Quality Assurance Authority for Education and Training (2009). Program Review Handbook, Vol. 2. Kingdom of Bahrain: Higher Education Review Unit.

Quality Assurance Authority for Education and Training (2012). Schools Review Unit: Review Framework and Guidance (2nd edition). Kingdom of Bahrain: Quality Assurance Authority for Education and Training.

Ramsden, P. (1991). A performance indicator of teaching quality in higher education: The Course Experience Questionnaire. Studies in Higher Education, 16(2), 129-150.

Ramsden, P. (1992). Learning to teach in higher education. London: Routledge.

Robinson, J., Reed, W., and Strauss, R. (2011). The impact of extended professional development and a comprehensive approach to assessment on teacher use of assessment for learning practices. Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, LA, April 8-12, 2011).

Ross, E. M., Ertmer, P. A., Johnson, T. E. (2001). Technology integration and innovative teaching practices: A staff development model for facilitating change. Paper presented at the National Convention of the Association for Educational Communications and Technology (24th, Atlanta, GA, November, 8-12, 2001).

(ERIC Documant Reproduction Service No. ED 470111).

Smith, M. L., Heinecke, W. & Noble, A. J. (1999). Assessment policy and political spectacle. Teachers College Record, 101(2), 157-191.

Stiggins, R. J. (2001). Student-Involved Classroom Assessment (3rd ed.). Upper Saddle River, NJ: Prentice-Hall, Inc.

Stufflebeam, D. L., & Shinkfield, A. J. (2007). Evaluation theory, models, and applications. San Francisco, CA: John Wiley.

Torrance, H. & Pryor, J. (1998). Investigating formative assessment: teaching, learning and assessment in the classroom. Buckingham: Open University Press.

University of Bahrain (2010). Institutional Review: Self-Evaluation Report. University of Bahrain: University of Bahrain Printing Press.

Vandeyar, S. & Killen, R. (2007). Educators' conceptions and practice of classroom assessment in post-apartheid South Africa. South African Journal of Education, 27 (5), 101-115.

Virginia Board of Education (2011). Guidelines for uniform performance standards and evaluation criteria for teachers. Richmond, VA: Author.

Wang, J. R., Kao, H. L., & Lin, S. W. (2010). Presevice teachers' initial conceptions about assessment of science learning: The coherence with their views of learning science. Teaching and Teacher Education, 26(3), 522-529.

Warren, E. & Nisbet, S. (1999). The relationship between the purported use of assessment techniques and beliefs about the uses of assessment, in: J. M. Truran & K. M. Truran (Eds.), 22nd Annual Conference of the Mathematics Education and Research Group of Australasia (Adelaide, SA, MERGA), pp. 515-521.

Weber, R. (1985). Basic Content analysis. Beverly Hills: Sage.

Weurlander, M. & Stenfors-Hayes, T. (2008). Developing medical teachers' thinking and practice: impact of a staff development course. Higher Education Research and Development, 27(2), 143-153.

Wilson, K. L., Lizzio, A., & Ramsden, P. (1997). The development, validation and application of the Course Experience Questionnaire. Studies in Higher Education, 22, 33-53.

Winterbottom, M., Brindley, S., Taber, K. S., Fisher, L. G., Finney, J., & Riga, F. (2008). Conceptions of assessment: Trainee teachers' practice and values. Curriculum Journal, 19(3), 193-213.

Yao, Y. (2015). Teacher perceptions of classroom assessment: A focus group interview. SRATE Journal, 24(2), 51-58.

APPENDIX 1

The Interview Questions

1- What is evaluation from your point of view?

2- What is the relationship between measurement, assessment, and evaluation?

3- What are the aims of evaluation for the subject you are teaching in the school?

4- What are the objectives of assessment as seen from your schools' point of view?

5- What type of assessment methods and techniques you use to assess your students?

6- What is the focus and emphasis of your methods of assessment in the classroom?

7- What are you usually trying to assess in your exams, tests, quizzes, projects, etc.?

8- How do you achieve congruence between the test questions and lesson objectives?

9- What criteria and standards guide your assessment of student learning?

10- How do you usually inform your students about their progress in learning?