Assessment Techniques and Students' Higher-Order Thinking Skills

Yousef Abosalem

Department of Mathematics and Science, Preparatory Program, Khalifa University, Abu Dhabi, UAE

Email address:

Yousef.abosalem@kustar.ac.ae

Abstract: Improving students' higher-order thinking skills is a collective experience; one teacher of a specific subject cannot alone improve the higher-order thinking skills, and it is a collaborative process between all subjects' teachers and can be taught for all levels of studying (Lawson, 1993; Shellens, & Valcke, 2005). Moreover, Benjamin (2008) argues that these skills can be developed in a cumulative fashion as students' progress through their courses and subjects and other experiences they get from their institutions. As well, by including their subjects by problem solving, critical thinking and decision making activities will help students enhance their higher-order thinking skills. In this paper a mathematics test in fractions was constructed and analyzed for both grades 8 and 9 to make sure how teacher-made tests are constructed and how much of them agreed with the Bloom's Taxonomy levels. The test consists of five sections or content areas the test was analyzed according to the behavior matrix. The results showed that all test items measure the lower three levels in Bloom's taxonomy which agrees with Stiggins, R. J., Griswold, M. M., and Wikelund, K. R. (1989) results that most of teacher-made tests measure the lower levels in Bloom's taxonomy. Moreover, 57.14% of the test items are applications and 28.57% are recognition items. These numbers are consistent with Boyd (2008) study, which indicated that the majority of teachers' assessment items focused on the lower levels of Bloom's Taxonomy. Moreover, Boyd concluded that 87% of the teachers' items that have participated in this study used level 1 of the taxonomy in 2003- 2004, and this percentage increased to 86% in 2005-2006. These numbers reflect the tendency of the assessment methods used in schools to ask students to recall information or to do routine question, which will not help students in improving their higher-order thinking skills.

Keywords: Assessment, Higher Order Critical Thinking, Mathematics

1. Introduction

Although many researchers (see e.g. Marzano, 1993; Ennis, 1993) have discussed and investigated higher-order thinking skills broadly, it has been misunderstood. Many researchers and educators considered higher-order thinking the same as the complexity of the questions raised or given to the learners. Complexity might be one of the aspects of higher-order skills, but it is not the only aspect.

Teachers in order to improve and develop students' higher-order thinking skills utilize and use different strategies. However, teachers must have knowledge of the specific skill of thinking. Improving these skills requires a great cooperation between all teachers of different subjects in different levels of studying to work together to achieve that goal. Moreover, cooperative learning will enhance students' thinking abilities. Through cooperative learning each student will articulate and share his ideas with other students who are involved in an interactive approach (Fogarty & McTighe, 1993) and consequently transfer these skills and apply them to other situations. Many researches (see e.g. Newman, 1990) argue that both lower and higher order thinking skills may be interwoven in the classroom, and the use of them depends on the nature of the student and the subject.

Developing students' skills requires creating assessment techniques that have abilities to help teachers in their job and reveal students' skills. Moreover, teachers are supposed to implement varieties of assessment methods such as performance- based assessment and stay away from the tests that require recalling knowledge such as observations, short answer questions and multiple-choice question which are most frequently used by class teachers (Doganay & Bal, 2010).

2. Learning Theories

In designing assessment techniques that are to be employed in classes requires giving our attention to the processes of student expected learning. There are many learning theories that deal with the issue of how people learn; three major learning theories tried to explain how students learn and acquire knowledge. These theories are: behaviorist, cognitive and constructivist, each of which has its own assumptions. Behaviorist considers learning as a resulting change in behavior and is seen to be linear and sequential, complex skills can be broken down into simple skills, each of which can be learned and mastered separately. Whereas, the cognitive theory perceives learning as taking place when the knowledge is internally contained and personalized. In contrast, the constructivist approach to learning is different from the previous two theories. This theory considers leaning takes place if the student is actively involved, participating and constructing new knowledge to be built on pre-existing knowledge, and the teacher's role is a facilitator (Anderson & Elloumi, 2004).

The development of students' higher order thinking is considered a central goal for all educators and educational stakeholders. They are trying to achieve that goal at all educational level. As well, it is considered a tool to develop the individuals and the community at the same time. Since the 1980s and 1990s, the attention increased towards conducting researches that aimed at how to improve the students' higher-order thinking skills. (Beyer, 1983; Costa, 1981; Sterberg, 1984). These studies showed the need for developing the teaching-learning process to improve these skills. Developing the students' higher-order skills can be achieved by two methods; 1) through lessons and special workshops in developing higher skills, and 2) through the regular mathematics classes and other school subjects (Beyr, 1983, NCTM, 1999). Moreover, Bereiter and Scardamalia (1987) argue that improving students' higher-order thinking skills can be achieved by constructing new models of curriculum and instruction techniques that can help in using critical thinking and problem solving approaches.

3. Definition of TERMS

3.1. Skill

Different definitions exist for skill. According to Ennis (1993), skill can be defined as sophisticated useful activity that requires intended training and organized practices. Others (Calfee, 1994) defined skill as efficiency and quality in performance, whether if it is defined as efficiency or quality in performance, skill is indicated to be learned or acquired behavior. This behavior has to be directed towards a specific purpose, has to be organized, and leads to achieving that purpose in the shortest time possible.

3.2. Critical Thinking

The word thinking refers to many different patterns of behavior, so it is difficult to define or choose a specific definition that includes the nature, means and products of thinking (Crenshaw, P., Hale, E., & Harper, S. L., 2011). However, Dewey (1966) considers thinking as a mental activity employed in different senses. Nosich (2012) stated that critical thinking consists several pertinent types; being reflective, involving standards, being authentic, and being reasonable. According to Facione (2011), critical thinking is "the process of reasoned judgment" (p. 3). Scriven and Paul (2004), of The Foundation of Critical Thinking, offered this definition: "Critical thinking is that mode of thinking - about any subject, content, or problem in which the thinker improves the quality of his or her thinking by skillfully taking charge of the structures inherent in thinking and imposing intellectual standards upon them" (para. 10). In this paper thinking will be defined according to Mosely et al. (2005) that "a consciously goal-directed process" (p. 12). Also, critical thinking will be referred to the evaluating process and assessing students' level of thinking.

3.3. Higher Order Thinking Skills

Lewis and Smith (1993) both are wondering if there is a difference between lower-order and higher-order thinking skills. In fact, the term "higher order" thinking skills seems a misnomer in that it implies that there is another set of "lower order" skills that need to come first. Newman (1990), in order to differentiate between the two categories of skills, concludes that the lower skills require simple applications and routine steps. In contrast and according to Newman (1993) higher order thinking skills "challenge students to interpret, analyze, or manipulate information" (P.44). However, Newman argues that the terms higher and lower skills is relative, a specific subject might demand higher skills for a particular student, whereas, another one requires lower skills. Splitting thinking skills into two categories will help educators in developing activities that can be done by slow learners before they can move to skills that are more sophisticated. As well as, to develop activities that can be performed by fast learners and place them in their appropriate level. Furthermore, this splitting helps educators in constructing remediation programs for slow learners consisting of drill and practice. By a process of remediation through repetition, students are expected to master the lower-order level thinking skills, which will help them in further stages to master the higher order skills. Moreover, by breaking down skills into simple skills and higher level skill will help curricula developer to design the subject's contents according to this splitting by focusing on basic skills in lower grades and in later grades, they can build the students' competences and higher-order thinking skills.

Educators consider higher-order thinking skills as high order thinking that occurs when the student obtains new knowledge and stores it in his memory, then this knowledge is correlates, organized, or evaluated to achieve a specific purpose. These skills have to include sub-skills such as analysis, synthesis and evaluation, which are the highest levels in Bloom's cognitive taxonomy.

A very important questions raised themselves here, why as educators are interested in developing the students' higher thinking skills? And why Bloom's cognitive taxonomy?

Ennis and Wheary (1995) answered the first question by stating the need to improve the students' higher thinking skills because developing these skills will diagnose the students' higher thinking levels, provide students with feedback about their levels of thinking and encourage them to think in a better way, provide teachers with information as to the extent they achieved the educational purposes, conducting studies on how to teach higher-order thinking skills. The second question will be answered later in this paper.

3.4. Assessment and Higher-Order Thinking Skills

Assessment is considered one of the challenging areas in educational theories and practices. It used to achieve a range of purposes by using different methods and techniques; each method has its own characteristics and properties. It can be used as a basis for reporting a particular student's performance as well as to evaluate the performance of the entire system. Moreover, assessment in mathematics can be used "to provide educators the opportunity to gain useful insight into students' understanding and knowledge of a specific subject, rather than just identifying their ability to use specific skills and apply routine procedures." (NCTM, 1995, p.87).

According to Airasian (1994) and Pellegrino, Chudowsky and Glaser (2001), assessment has three main purposes: to assist learning, to measure a particular student's achievement and to evaluate the whole program. So that, without good assessment techniques it is difficult to ascertain whether reforms in instruction and curriculum are working. The suitable assessment is one that can be used or leads to improvement in student's learning. Moreover, it can reveal the student's weakness and strength areas; the strength area to be enhanced and the weakness area to be treated.

A great shifting is occurring in education as researchers change their beliefs about teaching-learning processes. Many studies have been conducted to clarify how children learn and get their knowledge. The basic idea in most of these researches is that children are active builders of their knowledge, not just receptacles for knowledge. These studies created a great pressure on traditional teaching methods, traditional curricula, and consequently testing or assessment techniques. Furthermore, they clarify the relationship between achieving instructional goals and assessment. Bol and Strage (1993) both argued that there is a misalignment between higher-order thinking skills, instructional goals and types of test items used to measure these skills. In other words, there is no matching between the teachers' desire to achieve in their courses and the kind of assessment practices and test items students encounter in their courses. Along with that, teacher-made tests failed to reflect teachers' declared objectives (Haertel, 1991). Therefore, both test's and curricula designers have to work together to make sure that instruction and assessment take place at the same time in

order to enhance students' higher-order thinking skills by developing their abilities of communication, reasoning, and problem solving (NCTM, 1995).

In the past few decades, there has been a demand for better methods of assessing students' achievements in order to measure *what students can do with what they know*, rather than simply finding out *what they know* (Struyven, K., Dochy, F., Janssens, S., Schelfhout, W., & Gielen, S. 2006; Aschbaker, 1991; Linn, Baker, & Dunbar, 1991; Bracey, 1987; Boyd, 2008) as well as to fulfil the great demands for educators and policy makers for tests that reflects and measures the students' learning. Many educators believe that in order to teach higher-order thinking skills, to fill the gap between the teachers' assessment practices and instructional tasks or goals, and to implement new assessment ideas and classroom practices, a great change from traditional assessment which assess students' abilities to remember the facts (NRC, 2000), into authentic assessment that has the ability to reflect and measure the actual learning-teaching outcomes, and to evaluate and reform the goal of the new curricula and teaching strategies used in classes is required. As a result of this demand, other forms of assessment have been sought and many alternatives have been implemented. One of these forms is authentic assessment, which falls under the category of alternative assessment. It comprises the assessment of traditional academic content in combination with the skills and knowledge essential for lifelong learning. This type of assessment implies the usage of various techniques such as real-world situations (Mayer, 1992). The purpose, the context, the audience and the constraints of the given test, therefore, should all be related to real-world problems and or situations. These authentic forms of assessment have progressively replaced traditional types of assessment, such as paper and pencil tests and multiple-choice questions. Wiggins (1994) considered the paper and pencil tests as being invalid. Because the verbal ability rather than, or in addition to, the target ability is being tested, these new forms put a lot of emphasis on knowledge integration and the use of competencies in problem solving, and they help prevent misclassification of students who tend to perform relatively poorly on multiple-choice tests. Also, these tests require students to construct responses orally or in writing to a wide range of problems, create a product, or demonstrate application of knowledge in an authentic context (Calfee, 1994).

This emerging use of complex tasks and performance-based assessments has changed to a great extent the way teachers use assessment. The call for better forms has given birth to numerous questions concerning their disadvantages and relative benefits as compared to the simplest forms of assessment. However, broad comparisons are limited by the diversity of these forms, especially when taking into consideration the fact that each of these forms has its own benefits, disadvantages and issues involved.

In changing process from traditional assessment to the authentic assessment is faced with many challenges and obstacles. One of these challenges is the dominance of

standardized norm referenced multiple-choice tests, which are used most often to evaluate a lower thinking skills and educational achievement.

3.5. Assessment Methods

Traditionally, the term "assessment" refers to the process of gathering information in order to make evaluative decisions (Appi, 2000; Penta & Hudson, 1999), and was used in relation to quizzes and tests. Yet in a broader sense of the word, assessment is more about learning and teaching than just about testing and assessing student knowledge. As was already mentioned, new assessment methods and models (e.g. performance-based or alternative ones) are designed to introduce a wide range of opportunities and potential measures for students, with the objective to create and demonstrate what the students are able to do with their education programme.

Assessment models still place more emphasis on the student's performance than on the student's ability to use the obtained knowledge within specific terms taken out of the educational context. It should also be noted that assessment is an ongoing process as it is being conducted continually in various forms, providing teachers with a so-called "picture album" of the student's ability instead of the random and more isolated "snapshot" of the student's knowledge provided by traditional testing.

The traditional test is usually a one-time measure and is based on the achievement made by a given student on a particular day. Traditional assessments usually rely on a student's single correct answer per specific question (Wraga, 1994), usually omitting the student's demonstration of overall knowledge and their thought process. However, traditional test methods are still used in assessment modules, although they usually need to be combined with the ongoing assessment techniques in order to measure the performance of a particular student at a particular time, as well as the progress made by this student since the previous test in order to provide students and parents with useful feedback regarding how well the student is building important skills and knowledge (Wolf, 2007).

On the other hand, traditional tests are ineffective in measuring higher-order thinking skills or their abilities to deal with new and unusual problems and they give the impression that answers are always either right or wrong, and encourage memorization rather than understanding.

The differences between traditional and authentic assessments are as follows; traditional assessment usually implies a test on material that has been taught; this test usually covers limited educational material. Traditional assessment also implies testing only discrete measureable behaviours that focus on the *products* of learning rather than the *process* of learning (Bol, 1998). They also look for "right or wrong" responses which are more important than justifying one's methods and results; they assess for "coverage"-many items can be administered in a short period (Bennet, R. E, Morley, M., and Quardt, D., 2000), are designed to audit performance, and the questions must be

unknown to the students in order to ensure the validity of such a test.

In contrast to traditional assessments, authentic assessments are used to assess what the students can do with the educational material they have learned, and they aim to evaluate students' abilities in the real-world context (Authentic Assement Overview, 2001; Linda, B., Patricia, L. S., and O'Connel, A. A., 1998). Authentic forms of assessments encourage students to use their knowledge creatively, and challenge them to express their own interpretations of the material they have learned in class. Unlike traditional assessment, authentic assessment evaluates the accuracy with which a student is able to carry out a function within a given context, and assesses acquired knowledge (Stiggins, 1997; Viechniki, K. J., Barbour, N., Shaklee, B., Rohrer, J. & Ambrose, R., 1993; Sambell & McDowell, 1998).

Authentic assessments are competency-based, designed to elicit sophisticated versus naïve answers and assess for "un-coverage". They also aim to improve a student's performance instead of auditing it; and therefore, to be most effective, they should be known to students as much as possible in advance.

Similar to traditional paper and pencil tests, performance-based assessments are also important ways for students to demonstrate their talents to make connections and apply their knowledge, understanding, and higher-order thinking skills. However, while paper and pencil tests are usually relatively short, performance-based assessments may range from tasks that require a few days to be completed to huge projects that may take up to several weeks to complete (Eisner, 1999). As already mentioned, these types of tests are often referred to as authentic assessment, as they mirror expectations the students will encounter as adults. These assessments test higher-order thinking skills and require students to take an active part, unlike paper and pencil assessments where the students are passive test-takers.

Performance-based assessments should also be accessible to students with different experiences, learning styles, backgrounds, and abilities. It should also be noted that unlike traditional paper and pencil tests, performance-based assessments bear more resemblance to learning activities. However, they are different from them in two important ways, namely:

1) In performance-based assessments, the tasks should clearly and explicitly assess the targets which are being measured by the teacher (Doyle, 1983) or "the degree to which evidence and theory support the interpretations of test scores entailed by proposed uses of tests"(Miller & Linn, 2000, p. 367). In other words, they should be valid.

2) In these tests, the tasks should also have understandable and clear criteria for scoring, allowing the teacher to evaluate the results objectively, fairly, and consistently. In other words, this means that the tasks should be reliable (Williams & Ryan, 2000).

In addition, in these kinds of tests, the students have an option to participate in the process (for example, to define

scoring rubrics, or clarify the performance criteria); while in traditional paper and pencil assessments, the students simply provide responses.

Although some teachers consider the attributes of performance-based tests as disadvantages, overall, they have a positive impact on the diagnosis of students' errors and formative assessment (Khattri, et al., 1998). When the students are asked to perform authentic tasks rather than just selecting responses, they have the opportunity to demonstrate proficiency by doing something chosen by them. Moreover, authentic assessment gives the students the chance to approach a mathematical problem, for example, from several directions, which will enable teachers to find out their strengths and weaknesses, and consequently assess their performance abilities more accurately.

Some points to be taken into consideration are that performance-based assessments usually need much more time to create and administer than traditional paper and pencil assessments, and the scoring of these tasks may be more subjective, if compared to paper and pencil tests. This aspect becomes even more important if, in performance-based tasks, the teachers are not very explicit concerning the standards. Moreover, may be less careful in their attempt to define appropriate criteria for different levels of students' achievement (Stiggins et al, 1989).

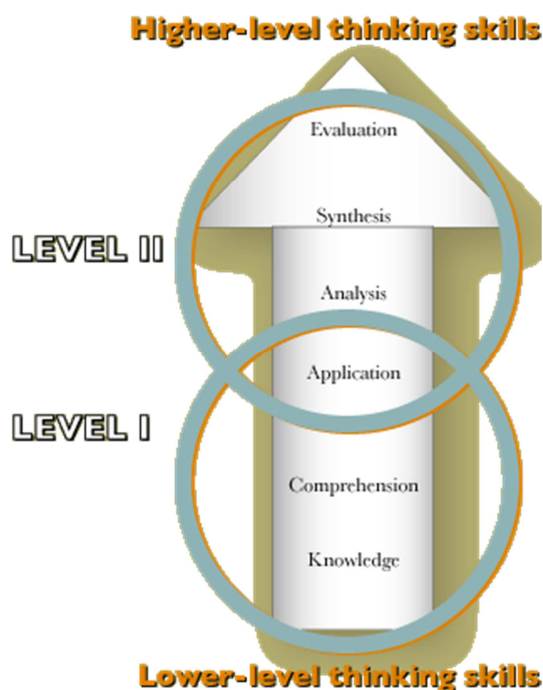## 4. Classroom Assessment of Higher Order Thinking Skills

Fig. 1. Bloom's Taxonomy (Bloom, 1956 in USF).

Critical thinking can be developed in both traditional and online classrooms. Incorporating critical thinking in the classroom allows participants to question assumptions, identify bias and engage in rigorous and self-disciplined discussion. Recently, most attempts to integrate cognitive skills into test construction process were guided by Bloom's Taxonomy of cognitive domain (Figure-1) to classify the cognitive skills students have to gain and acquire. Constructing tests according to this behavioral approach requires organization according to content by behavior the matrix. This matrix shows the content area that's supposed to be tested and the skills test takers are expected to show in each content area.

Many researchers revealed extensive use of test items at Bloom's lowest level (Crooks & Collins, 1986; Fleming & Chambers, 1983; Haertel, 1986; Stiggins, et al., 1989). For example, Fleming and Chambers (1983) (cited in Crooks & Collins, 1986) by analyzing 8800 tests, found that that 80% of the tests that were analyzed were at knowledge level of Bloom's taxonomy. Stiggins, et al (1989) outlined that the assessments based on teachers' observations and judgments, and tests constructed by them are not commonly suitable to measure high-order thinking skills. Teacher-made tests overly emphasize memory for procedures and facts (Porter, A. C., Kirst, M. W., Osthoff, E. J., Smithson, J. S., & Schneider, S. A., 1993; Burns, 1985) and tend to give more weight to lower levels than the teachers' declared objectives would justify. Few items of their tests were prepared to measure skills above the first three levels of Bloom's Taxonomy of knowledge, comprehension, and application and the majority of them are recall items. Furthermore, they concluded that 55% of the test items used by 36 teachers who taught mathematics, science, social studies and language art at grade 1-12 were recall measures, 19% of them were inference, 16% analysis, 5% comparison and 5% evaluation. From the percentages mentioned above, it can be concluded that the majority of teachers' tests are not constructed to measure the higher order skills. This might be due to the fact that many teachers are not trained in how to construct this type of questions or they are reluctant to use a new testing approach such authentic assessment in their classes due to the time needed to design such tests. Moreover, Boud (2008) found that teachers were unable to construct tests that provide insight into students' thinking. In another study conducted by Bol and Strage (1993) to find out the relationships among teachers' assessment practices and their student outcome and study skill development goals, outlined that 53% of test items required only basic knowledge, while almost none required application, nearly two-third (65%) of test items were recognition items. The teachers' justification is that using questions to assess higher order thinking skills will confuse their students, increase the anxiety level and number of failures (Doyle, 1983). However, the use of higher-level questions which require the student to integrate and use different ideas levels ranging from simple to sophisticated ideas will improve students' learning which is considered as the process of acquiring knowledge or skills or attitudes towards subjects which consequently involves changes in behavior (Theresa, 2015). Also, learning is a tool to transfer knowledge to new situations which is considered a function of the relationship between what is learned and what is tested

and development of learning skills (NRC, 2000). Whereas, Lawson (1993) argues that in order to improve higher-order thinking skills, the teachers have to encounter their students with situations in which they struggle to answer provoked questions then to reflect on these answers and on the method of getting them.

5. Mathematics Test Analysis

In order to improve our schools, new assessment techniques have to be employed. Assessment has two purposes; help stakeholders to make their decisions in issues that related to students and schools. Also, it serves as a learning motivation tool. To achieve the second purpose, schools have to use and implement high- stakes testing. However, high-stakes testing supposed to be designed to fit all students' academic levels.

To analyze the mathematics test, the matrix behavior was used. This matrix is used by educators to separate the test's content and behavior in terms of Bloom's taxonomy. Also, the matrix is considered as a planning tool that gives instructors a clear idea on their teaching and areas that need to be emphasized.

In order to clarify the matrix behavior and to make sure how teacher-made tests are constructed, and how much of them agreed with the Bloom's Taxonomy levels, a mathematics test were constructed for both grades 8 and 9

(Appendices A & B). The test was selected from a public school in Abu-Dhabi and was developed by mathematics teachers in that school. In order to attain the test technical qualities, it was given for mathematics experts to check its face validity and its visibility. With respect to the test reliability, it was conducted to two sections (one in grade 8 and one in grade 9) about 50 students in both grades. By using Cronbach Alpha, the reliability coefficient was 0.87 which is considered reliable with respect of many psychometricians. The test consists of five sections or content areas as shown in table-1. The test was analyzed according to the behavior matrix. Table-1 showed that all test items measure the lower three levels in Bloom's taxonomy which agrees with Stiggins et al (1989) results that most of teacher-made tests measure the lower levels in Bloom's taxonomy. Moreover, 57.14% of the test items are applications and 28.57% are recognition items. These numbers are consistent with Boyd (2008) study, which indicated that the majority of teachers' assessment items focused on the lower levels of Bloom's Taxonomy. Moreover, Boyd concluded that 87% of the teachers' items that have participated in this study used level 1 of the taxonomy in 2003- 2004, and this percentage increased to 86% in 2005-2006. These numbers reflect the tendency of the assessment methods used in schools to ask students to recall information or to do routine question, which will not help students in improving their higher-order thinking skills.

Table 1. *Shows the behavior matrix (Blue print) for the mathematics test.*

Content Areas								
Cognitive Levels		**Fractions' Concepts**	**Adding & Subtracting fractions**	**Multiplying & Dividing fractions**	**Combined operations with fractions**	**Word problems with fractions**	**Total**	**Percentages**
Behavior	Evaluation	0	0	0	0	0	0	0%
	Synthesis	0	0	0	0	0	0	0%
	Analysis	0	0	0	0	0	0	0%
	Application	3	5	3	4	3	18	57.14%
	Comprehension	5	1	3	0	1	10	14.29%
	Knowledge	7	0	0	0	0	7	28.57%
	Total	15	6	6	4	4	35	
	Percentages	42.90%	17.14%	17.14%	11.41%	11.41		100%

6. Conclusion

Lewis and Smith (1993) consider critical thinking, problem solving, decision making and creative thinking aspects of higher-order thinking skills. It can be measured by different assessment methods such as performance tests, portfolios, projects and multiple-choice items with written justification (Ennis, 1993). Therefore, teacher-made test, which they are used commonly in our schools, are not suitable to measure higher-order thinking skills. It might be because many teachers are not trained to prepare different types of assessments. Whereas, many researchers (Shepard, 1989; Wiggins, 1989, 1993; Paul & Nosich, 1992) proposed

performance based assessment as a tool to measure higher-order thinking skills.

Improving students' higher-order thinking skills is a collective experience; one teacher of a specific subject cannot alone improve the higher-order thinking skills, and it is a collaborative process between all subjects' teachers and can be taught for all levels of studying (Lawson, 1993; Shellens, & Valcke, 2005). Moreover, Benjamin (2008) argues that these skills can be developed in a cumulative fashion as students' progress through their courses and subjects and other experiences they get from their institutions. As well, by including their subjects by problem solving, critical thinking and decision making activities will help students enhance their higher-order thinking skills. Moreover, by helping students to

be critical thinkers we enable "them to make informed decisions about social issues" (Trevino, 2008, p. 14).

In conclusion, it may be said that although traditional paper and pencil tests are still effective for assessing some student skills such as listening, because they can be measured, and the measurements can be used to compare students, schools, districts, states and countries. However, they fail to assess student's productive skills (e.g. writing or speaking) and to prepare students for the real world. Paper and pencil tests (e.g., completion, true-false, matching, multiple choices, short answer, essay, etc.) imply that students are passive test takers. Moreover, many studies outlined that these types of tests tighten the content that teachers teach in their classes (Abrams et al., 2003; Firestone et al., 1998) and do not provide teachers with a clear idea into student actual learning (Black & William, 1998). While performance-based assessments such as rubrics, checklists, portfolios, and reflections, encourage the student to display, his or her best work because they are designed to promote students' participation in the tests. Moreover, it provides educators with an accurate picture of what students know (knowledge). However, Benjamin (2008) argues that some of the performance assessment methods such as portfolio are not suitable for assessing such kind of skills because portfolio assessment suffers from serious reliability and validity problems, instead he proposed comparative tests to do that such as the College Assessment of Academic Progress(CAAP).

Finally, performance-based assessments afford us with information about the students' daily improvement, and insight into the process of learning, because this type of assessment requires students to demonstrate that they have mastered specific skills and competencies by performing or producing something. Moreover, using many sources of assessments gives the teachers a comprehensive view of student progress, and can help them gain an understanding of how students think and learn new skills. As well as, assessment should be more than an event inserted at the end of a learning period (Shepard, 2000). So that, I do believe that using this approach of assessment will be helpful in improving and evaluating students' higher-order thinking skills, rather than using one shot tests such as multiple-choice tests.

While we are excited about implementing alternative assessment, other issues have to considered: time constraints, subjectivity, validity, economic issues, good training in assessment, cultural bias, and the need for extensive teachers' professional development.

Appendix (A)

Mathematics Test

Name	School
Date	

This test is designed to explore patterns of errors in the four operations made by grade 8 and grade 9 students on fractions.

This test will evaluate your knowledge of fractions and diagnosis the weak areas. Your result on this test will not effect on your mathematics score in school.

Instructions:-

1. Write your name and the name of your school before start answering this test.
2. Calculators are not permitted to be used in any part of this test.
3. Show your work when it is required.
4. Scrape paper is not permitted; use the pack side of the test booklet.
5. All work should be done in pencil.
6. This test contains five sections, with 35 questions. You have to solve all questions in this booklet
7. Answer the multiple choice questions on the provided answer sheet at the end of this booklet.
8. Multiple-choice questions, one point each, three points for problems.

Section (1): Concepts.

For each of the following questions (1-15), select the correct answer.

1) A fraction is:
   Way to show division with the part over the whole.
   a) A whole number and a fraction.
   b) The top number of a fraction.
   c) A whole number.
2) The number of parts there are in the whole is called the:
   a) Numerator   b) Denominator
   b) Mixed number   d) Whole number
3) The top number of a fraction is called:
   a) Numerator   b) Denominator
   c) Mixed number   d) Whole number
4) Mixed number is:
   a) The top number of a fraction.
   b) The bottom number of a fraction.
   c) A mixture of whole number and proper fraction.
   d) A mixture of whole number and improper fraction.
5) Equivalence fractions are:
   a) Fractions with the same numerators.
   b) Fractions with the same denominators.
   c) Fractions have the same value.
   d) Two different amounts.
6) Which is an improper fraction?
   a) 4   b) 8
   c) 101   d) 6
7) The least common denominator (LCD) of $\frac{5}{6}, \frac{3}{8}, \frac{7}{12}$ is:
   a) 96   b) 48
   c) 26   d) 36
8) $\frac{1}{5}$ and $\frac{1}{9}$ are:
   a) Mixed Numbers.   b) Equivalent Fractions.
   c) Proper Fractions.   d) Improper Fractions.
9) Which of the following fractions is in reduced form (Simplest form)?

a) $\dfrac{3}{96}$ b) $\dfrac{11}{44}$

c) $\dfrac{5}{16}$ d) $\dfrac{16}{42}$

10) $6\dfrac{3}{10}$ Is the same as?

a) $\dfrac{60}{3}$ b) $\dfrac{61}{3}$

c) $\dfrac{63}{10}$ d) $\dfrac{57}{10}$

11) $\dfrac{37}{5}$ Is the same as?

a) $7\dfrac{2}{5}$ b) $7\dfrac{5}{2}$

c) $37\dfrac{1}{5}$ d) $32\dfrac{1}{5}$

12) The reciprocal of $\dfrac{3}{5}$ is:

a) $\dfrac{3}{3}$ b) $\dfrac{5}{3}$

c) $\dfrac{10}{6}$ d) $\dfrac{10}{3}$

13) Which expression is correct?

a) $\dfrac{3}{4} > \dfrac{6}{8}$ b) $\dfrac{7}{8} < \dfrac{1}{3}$

c) $\dfrac{3}{5} = \dfrac{18}{30}$ d) $\dfrac{4}{5} = \dfrac{12}{20}$

14) Which is an improper fraction?

a) $\dfrac{1}{9}$ b) $\dfrac{15}{11}$

c) $\dfrac{100}{102}$ d) $\dfrac{3}{6}$

15) Which one is Not an example of using fraction in life?

a) 4 people in a family.   b) Size $10\dfrac{1}{2}$ shoes.

c) $4\dfrac{1}{3}$ Yards of material.   d) $\dfrac{1}{2}$ Teaspoon sugar.

Section (2): Addition & Subtraction. (Show your work).

Add or subtract each pair of fractions below. Make sure your answers in simplest forms.

1) $3\dfrac{2}{5} + 4\dfrac{1}{3} =$ 2) $3\dfrac{5}{8} - 1\dfrac{1}{4} =$

3) $14\dfrac{3}{8} - 4 =$ 4) $8 - 2\dfrac{5}{9} =$

5) $7\dfrac{1}{2} - 2\dfrac{4}{5} =$ 6) $9\dfrac{1}{3} - 4\dfrac{5}{9} =$

Section (3): Multiplication & Division. (Show your work).

Solve the following questions. Make sure your answers in simplest forms.

1) $\dfrac{5}{9} \times \dfrac{6}{25} =$ 2) $2\dfrac{1}{5} \times 7\dfrac{8}{11} =$

3) $\dfrac{3}{8} \times \dfrac{4}{9} \times \dfrac{5}{11} =$ 4) $2\dfrac{3}{4} \div 5\dfrac{5}{6} =$

5) $\dfrac{5}{8} \div 13\dfrac{1}{8} =$ 6) $1\dfrac{2}{7} \times \dfrac{2}{7} \times 2\dfrac{1}{2} =$

Section (4): Combined Operations. (Show your work).

Solve the following questions. Make sure your answers in simplest forms.

1) $2\dfrac{1}{4} - 1\dfrac{1}{3} + 2\dfrac{1}{6} =$

2) $9\dfrac{3}{4} - 4\dfrac{2}{3} + 1\dfrac{7}{12} - 2\dfrac{1}{6} =$

3) $1\dfrac{1}{2} - \dfrac{2}{5} \div \dfrac{1}{5} \times \left(\dfrac{3}{8} + \dfrac{1}{8}\right)^2 + \dfrac{2}{3} =$

4) $\dfrac{3}{4} + \dfrac{2}{3} \div \dfrac{1}{4} + \left(\dfrac{2}{3} \times \dfrac{3}{4} \div \dfrac{1}{2}\right) =$

Section (5): Word problems.

Solve the following word problems. Show your strategy in solving the problem.

1. John is reading a book that is 697 pages long. He tells a friend that he is about $\dfrac{3}{4}$ of the way done. About how many more pages must John read before he finishes the book?

2. Lyle tallied the number of vehicles that crossed Confederation Bridge in one hour. He found that $\dfrac{3}{4}$ of the vehicles were automobiles and $\dfrac{1}{5}$ were trucks. What fraction describes the number of vehicles that were not automobiles or trucks?

3. Dana had a piece of fabric 18 inches wide. She wanted to cut it into thin strips $\dfrac{3}{4}$ of an inch wide to use in here quilting project. How many strips she get from the piece of fabric?

4. Angela made her own powdered drink mix for summer coolers. She used $\dfrac{2}{3}$ cup Tang, $\dfrac{1}{2}$ cup sugar, and $\dfrac{3}{4}$ cup Crystal Light lemonade. How much mixture did she end up with for each batch?

The end

Appendix (B)

Test Table of specifications (Blue – Print)
Numerals in Parentheses Refer to Specific items on the Test

Contents	Knowledge	Comprehension	Application	Total	Percentages
Section (1) Concepts					
1. Fraction's definition					
2. Types of fractions (proper, Improper Fractions& mixed numbers)					
3. GCF & LCD	7 (1, 2, 3, 4, 5, 14, 15)	5 (6, 8, 10, 11, 12)	3 (7, 9, 13)	15	42.90%
4. Lowest terms(Simplest form)					
5. Equivalent fractions					
6. Comparing size of fractions					
Section (2) Adding & Subtracting fractions					
1. Adding fractions with like denominators					
2. Adding fractions with unlike denominators using (LCD)	0	1 (1)	5 (2, 3, 4, 5, 6)	6	17.14%
3. Subtracting fractions with like denominators					
4. Subtracting fractions with unlike denominators using (LCD)					
Section (3) Multiplying & Dividing fractions					
1. Multiplication of fractions	0	3 (2, 3, 5)	3 (1, 4, 6)	6	17.14%
2. Division of fractions					
Section (4) Combined operations with fractions Combined operations with fractions	0	0	4 (1, 2, 3, 4)	4	11.41%
Section (5) Word problems with fractions Word problems based on fractions	0	1 (1)	3 (2, 3, 4)	4	11.41%
Total	7	10	18	35	
Percentages	28.57%	14.29%	57.14%		100%

References

[1] Abrams, L. M., Pedulla, J. J., & Madaus, G. F. (2003). Views from the classroom: Teachers' opinions of statewide testing programs. *Theory into Practice, 42*(1)*, pp. 18-29.

[2] Airasian, P. W. (1994). *Classroom assessment.*2nd ed. McGraw Hill, New York.

[3] Anderson, T., and Elloumi, F. (2004). *Theory and Practice of Online Learning* [online]. Athabasca University [Accessed 11 May 2010]. Available at: http://cde.athabascau.ca/online_book.

[4] Appl, D. J. (2000). Clarifying the preschool assessment process: Traditional practices and alternative approaches. *Early Childhood Education Journal*, 27(4), pp. 219-225.

[5] Aschbaker, P. R. (1991). Performance assessment: State activity, interest, and concerns. *Applied Measurement in Education,* 4(4), pp. 275–288.

[6] *Authentic Assessment Overview*. (2001) Person development group. [Online]. [Accessed 11 April 2010]. Available at: www.ithaca.edu/jwiggles/t&m/check.htm.

[7] Benjamin, R. (2008). The Case for Comparative Institutional Assessment of Higher-Order Thinking Skills. *Change*, 40(6), pp. 51-55.

[8] Bennet, R. E, Morley, M., and Quardt, D. (2000). Three response types for broadening the connection of mathematical problem solving in computerized tests. *Applied Psychological Measurement, 24*(4), pp. 294-309.

[9] Bereiter, C. and Scardamalia, M. (1987). An attainable version of high literacy: Approaches to teaching higher-order thinking skills in reading and writing. *Curriculum Inquiry*, 17, pp. 9-30.

[10] Beyer, B. (1983). Common sense about teaching thinking. *Educational Leadership*, 41(3), pp. 44-49.

[11] Bol, L., & Strage, A. (1993). Relationships among teachers' assessment practices and their student outcome and study skill development goals. (*ERIC Document Reproduction Service No. ED 367 639*) [online]. [Accessed 05 May 2010] available at: http://www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/15/5b/34.pdf.

[12] Bol, L. (1998). Influence of experience. Grade level and subject area on teachers' assessment practices. *The Journal of Educational Research*, 91(6), pp. 323-330.

[13] Boyd, B. (2008). Effects of state tests on classroom test items in mathematics, *School Science and Mathematics*, 108(6), pp. 251-261.

[14] Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan,* 80(2), pp. 139-149.

[15] Bloom, B. S. (Ed.). (1956). A learning in classroom instruction. New (1975). Masteri, objectives. Handbook I, the taxonomy of educational domain. New York: Longman.

[16] Bracey, G. W. (1987). Measurement-driven instruction: Catchy phrase, dangerous practice, Phi Delta Kappan, 68(9), pp. 683-686.

[17] Burns, M. (1985). The role of questioning. *The Arithmetic Teacher*, 32(6), pp. 14-17.

[18] Calfee, R. C. (1994). Cognitive assessment of classroom learning. *Education and Urban Society*, *26*(4), pp. 340-351.

[19] Christine Sereni (2015). Teaching Strategies for Critical Thinking Skills. Academic Exchange Quarterly Fall 2015 ISSN 1096-1453, 19(3).

[20] Costa, A. (1981). Teaching for intelligent behavior. *Educational Leadership*, 39(1), pp. 29-32.

[21] Crenshaw, P., Hale, E., & Harper, S. L. (2011). Producing intellectual labor in the classroom: The utilization of a critical thinking model to help students take command of their thinking. Journal of College Teaching & Learning, 8(7), 13-26. Retrieved from http://search.ebscohost.com/.

[22] Crooks, T. J. (1888). The Impact of Classroom Evaluation Practices on Students. *Review of Educational Research*, 58(4), pp. 438-481.

[23] Dewey, J. (1966). Democracy and education: An introduction to the philosophy of education. New York: Collier-Macmillan.

[24] Doyle, W. (1983). Academic work. *Review of Educational Research*, 53, pp. 159-199.

[25] Doganay, A. and Bal, A. P. (2010). The Measurement of Students' Achievement in Teaching Primary School Fifth Year Mathematics Classes. Educational Science: *Theory & Practice*, 10(1), pp. 199-215.

[26] Eisner, E. (1999). The uses and limits of performance assessment. *Phi Delta Kappan*, 80(9*)*, pp. 658-660.

[27] Ennis, R. H. (1993). Critical Thinking Assessment. *Theory in Practice*, 32(3), pp. 179-186.

[28] Ennis, R. H. and Wheary, J. (1995). Gender Bias in Critical Thinking: Continuing the Dialogue. *Educational Theory*, 45(2), pp. 213-224.

[29] Facione, P. (2011). Think critically. Boston, MA: Prentice Hall

[30] Firestone, W. A., Mayrowetz, D., and Fairman, J. (1998). Performance-based assessment and instructional change: The effects of testing in Maine and Maryland. *Educational Evaluation and Policy Analysis,* 20(2), pp. 95-113.

[31] Fleming, M. & Chambers, B. (1983). Teacher-made tests: Windows on the classroom. In W. E. Hathaway (Ed.), *Testing in the schools: New directions for testing and measurement,* pp. 29-38. San Francisco, CA: Jossey-Bass.

[32] Fogarty, R. and McTighe, M. (1993). Educating Teachers for Higher Order Thinking: The three-Story Intellect. *Theory into Practice*, 32(3), pp. 161-169.

[33] Gronlund, N. E. (1998) *Assessment of student achievement* (8th ed.) Boston, Pearson/Allyn and Bacon.

[34] Hollander, S. K. (1978). A literature review: Thought processes employed in the solution of verbal arithmetic problems. *School Science and Mathematics*, *78*, pp. 327-335.

[35] Haertel, E. H. (1991). New forms of teacher assessment. *Review of Research in Education*, 17(1), pp. 3-29.

[36] Kalyuga, S. (2006). Rapid assessment of learners' proficiency: A cognitive load approach. *Educational Psychology*, 26(6), pp. 735-749.

[37] Khattri, N., Reeve, A., and Kane, M. (1998). *Principles and practices of performance assessment.* Mahwah, NJ: Lawrence Erlbaum.

[38] Lai, A. F. (2007). The development of computerized two-tier diagnostic test and remedial learning system for elementary science learning. Seventh IEEE International Conference on Advanced Learning Technologies (ICALT). [Online]. [Accessed 29 April 2010].

[39] Available at: http://csdl2.computer.org/comp/proceedings/icalt/2007/2916/00/29160735.pdf.

[40] Lawson, A. (1993). At What Levels of Education is the Teaching of Thinking Effective? *Theory into Practice*, 32(3), pp. 170-178.

[41] Lewis, A. and Smith, D. (1993). Defining Higher Order Thinking. *Theory into Practice*, 32(3), pp. 131-137.

[42] Linda, B., Patricia, L. S., and O'Connel, A. A. (1998). Influence of experience, grade level, and subject area on teachers' assessment practices. *The Journal of Educational Research*, 91(6), pp. 323-331.

[43] Linn, R. L., Baker, E. L., & Dunbar, S. B. (1991). Complex Performance-based assessment: expectations and validation criteria. *Educational Researcher*, 20(8), pp. 15-21.

[44] Marzano, R. J. (1993) How classroom Teachers Approach the Teaching of Thinking, *Theory into Practice*, 32(3), 154-160.

[45] Mayer, C. (1992). what's the difference between authentic and performance assessment? *Educational Leadership*, 49(8), pp. 39-42.

[46] Michael Scriven Richard Paul (2004). Defining Critical Thinking. Retrieved from 10/9/2016 from http://www.criticalthinking.org/aboutCT/definingCT.shtml.

[47] Miller, M. D., & Linn, R. L. (2000). Validation of performance –based assessments. *Applied Psychological Measurement,* 24(4), pp. 367-378.

[48] National Research Council (NRC). (2000). *How People Learn: Brain, Mind, Experience, and School*: Expanded Edition, National Academy Press, Washington, D. C.

[49] National Council of Teachers of Mathematics. (1995). *Assessment standards for school mathematics*. Reston, VA: National Council of Teachers of Mathematics.

[50] National Council of Teacher of Mathematics (NCTM) (1999) *Developing Mathematical Reasoning in Grades K- 12* Yearbook of the National Council of Teacher of Mathematics, Reston. VA: National Council of Teachers of Mathematics.

[51] Newman, F. M.(1990) Higher Order Thinking in Teaching Social Studies: A rationale for the assessment of Classroom Thoughtfulness, Journal of Curriculum Studies, 22(1), pp. 41-56.

[52] Nosich, G. (2012). Learning to think things through. A Guide to critical thinking across the curriculum. (4th ed). Boston, MA: Pearson.

[53] Paul, R., & Nosich, R. (1992). *A model for the national assessment of higher order thinking*. (ERIC Document Reproduction Service No. ED 353 296).

[54] Pellegrino, J, Chudowsky, N & Glaser, R (eds) (2001). *Knowing what students know: The science and design of educational assessment: A report of the National Research Council,* National Academy Press, Washington DC.

[55] Penta, M. Q., & Hudson, M. B. (1999). Evaluating a practice-based model to develop successful alternative assessment at instructionally innovative elementary Schools. Paper presented at the annual meeting of the American Educational Research Association (AERA): (pp. 1-22). Montreal, Quebec, Canada. [Online]. [Accessed 4 May 2010]. Available at: http://eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storag e_01/0000019b/80/15/e8/6f.pdf.

[56] Porter, A. C., Kirst, M. W., Osthoff, E. J., Smithson, J. S., & Schneider, S. A. (1993). Reform up close: An analysis of high school mathematics and science classrooms (Final Report to the National Science Foundation on Grant No. SPA-8953446 to the Consortium for Policy Research in Education). Madison, WI: University of Wisconsin– Madison, Wisconsin Center for Education Research.

[57] Sambell, K., and McDowell, L. (1998). The construction of the hidden curriculum: messages and meanings in the assessment of student learning. *Assessment and Evaluation in Higher Education,* 23 (4), pp. 391-402.

[58] Shellens, T., & Valcke, M. (2005). Collaborative learning in asynchronous discussion groups: What about the impact on cognitive process? *Computers in Human Behavior,* 21(6), pp. 957-975.

[59] Shepard, L. A. (1989). Why we need better assessments. *Educational Leadership,* 46(7), pp. 4-9.

[60] Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher*, 29, pp. 4-14.

[61] Stiggins, R. J., Griswold, M. M., and Wikelund, K. R. (1989). Measuring thinking skills through classroom assessment. *Journal of Educational Measurement,* 26(3), pp. 233−246.

[62] Stiggins, R. J., Frisbie, D. A. and Griswold, P. A. (1989). Inside high school grading practices: Building a research agenda. Educational Measurement: Issues and Practices, 8(2), PP. 5-14.

[63] Stiggins, R. J. (2002). Assessment crisis: The absence of assessment for learning. *Phi Delta Kappan*, *83*(10), 758-765.

[64] Stiggins, R. J. (1999). Assessment, student confidence, and school success. *Phi Delta Kappan*, *81*(3), 191-198.

[65] Stiggins, R. J. (1997). *Student-centered classroom assessment.* Upper Saddle River, NJ: Merrill, an imprint of Prentice Hall.

[66] Struyven, K., Dochy, F., Janssens, S., Schelfhout, W., and Gielen, S. (2006). The overall effects of end-of-course assessment on student performance: A comparison between multiple-choice testing, peer assessments, case –based assessment and portfolio assessment. *Studies in Educational Evaluation*, 32(3), pp. 202-222.

[67] Sterberg, R. (1984). How can we teach intelligence? *Educational Leadership*, 42(1), pp. 38-48.

[68] Theresa Ebiere Dorgu. Different Teaching Methods: A Panacea for Effective Curriculum Implementation in the Classroom. International Journal of Secondary Education. Special Issue: Teaching Methods and Learning Styles in Education. Vol. 3, No. 6-1, 2015, pp. 77-87. doi: 10.11648/j.ijsedu.s.2015030601.13.

[69] Trevino, E. (2008, September 13). It's critical to learn how to be critical thinkers. El Paso Times, Retrieved from http://www.nwp.org/cs/public/print/resource/2821.

[70] University of South Florida (SFU) (2010). Classroom Assessment, [online]. [Accessed April 20, 2010]. Available at: http://fcit.usf.edu/assessment/selected/responsea.html.

[71] Wiggins, G. (1993). Assessment: Authenticity, context, and validity. *Phi Delta Kappan, 75,* 200−214.

[72] Wiggins, G (1989). 'A true test: Toward more authentic and equitable assessment', *Phi Delta Kappan,* 70(9), pp. 703–713.

[73] Viechniki, K. J., Barbour, N., Shaklee, B., Rohrer, J. & Ambrose, R. (1993). The impact of portfolio assessment on teacher classroom activities. *Journal of Teacher Education*, 44(5), 371-377.

[74] Wiggins, G. (1994). *Toward more authentic assessment of language performances*. In C. Hancock(Ed), *Teaching, testing, and assessment: Making the connection. Northeast conference reports.* Lincolnwood, IL: National Textbook Co.

[75] Williams, J. & Ryan, J. (2000). National testing and the improvement of classroom teaching: Can they coexist? *British Educational Research Journal*, 26(1), pp. 49-73.

[76] Wolf, P. J. (2007). Academic improvement through regular assessment. Peabody Journal of Education, 82(4), pp. 690–702.

[77] Wraga, W. G. (1994). Performance assessment: A golden opportunity to improve the future. *NASSP Bulletin*, 78(563), pp. 71-79.