Measurement and assessment for science education in the Turkish educational context: Problems and reflections

Mehmet KÜÇÜK

Karadeniz Technical University, Artvin Faculty of Education, Turkey

Salih ÇEPNİ

Karadeniz Technical University, Fatih Faculty of Education, Turkey

Email: mehmetkucuk@tused.org

Received 23 July, 2004
Revised 30 November, 2004

Contents

- Abstract
- 1. Introduction
  - 1.1. Need for Measurement and Assessment
  - 1.2. Overview to the Measurement and Assessment Area in the World
  - 1.3. Science Learning and Assessment
- 2. Measurements and Assessment in the Turkish Context
  - 2.1. Structure of the Turkish Educational System
  - 2.2. Review of the Literature on Measurement and Assessment Approaches in Turkey
- 3. Future Trends for Measurements and Assessment of Science Education
- References

Abstract

Research on teaching and learning supports the movement toward new assessment approaches. Educational assessment has an important effect on learners' growth, achievement and self-esteem. International literature includes a series of conceptual models that introduce the means and importance of student assessment from different perspectives. In this article, first we examined these new trends and focused on some new assessment methods in science education and then looked at the Turkish context based on the discussed criteria for science assessment. Then we analyzed some thoughts about the future of science assessment.

Key Words: Measurement and Assessment, Science Education, New Trends, Turkish Context.

Introduction

In the last few decades, some special value has been attached to student and school assessment approaches in educational literature (Amrein & Berliner, 2002; Doran, Tamir & Chan, 1995; Neill, 1997; Ulmer, 2001). National Research Council of the USA and some others have made an important contribution to the improvement and development of educational assessment area. In this article we first examined these new trends and focused on some new assessment methods in science education and then looked at the Turkish educational context based on the explained criteria for science assessments. Then we discussed some thoughts about the future of science assessment.

The USA National Research Council created its own National Science Education Standards (NRC, 1996). This document has been accepted to provide the best research-based thinking on science assessment. It includes the standards that have been developed for students to show how much science they know and how they know it as well as to start to have an understanding of various uses for science assessment. It has recommended the following shifts in assessment process (NRC, 1996):

- From what is easily measured to what is highly valued,
- From assessing knowledge to assess understanding and reasoning,

- From finding out what students do not know to finding out what they know,
- From the end of assessment term by teachers to ongoing assessment by students.

Looking at these shifts, it is obvious that learners are to be more active in learning and teaching contexts, teachers are to focus on making the hint points appear and to raise each student's own potential to the highest level. Just in this way, we can use the full potential of all learners. In addition, this document provided a shift toward performance based assessment science assessments in schools (NRC, 1996). This let evidence of higher learning in the form of demonstrated reasoning and problem solving skills which has been also supported in an intense value in the science education literature. These trends contributed to implementing some alternative and more authentic science assessments, such as those in which students can demonstrate what they know and can do. This will be explained in the context of a new concept of constructivist assessment through subsequent sessions.

Besides NRC, we also want to discuss a document called "Principles and Indicators for Student Assessment Systems" developed by National Forum on Assessment. In an article written by Neill (1997), he drew on principles to outline what a new assessment system would look like and to suggest some actions that can be taken into consideration to further assessment reform. The most important and attractive part of this document is that it requires a radical reconstruction of assessment practices in schools and suggests that student learning be made central to assessment reform. The seven principles on which all the participants of the forum agreed are given below:

- The primary purpose of assessment is to improve student learning;
- Assessment for other purposes supports student learning;
- Assessment systems are fair to all students;
- Professional collaboration and development support assessment;
- The broad community participates in assessment development;
- Communication about assessment is regular and clear;
- Assessment systems are regularly reviewed and improved.

Here, it is stressed that the aim of assessment is to improve and develop student learning not just to find out how good students are at some kinds of examinations, that assessment system should be convenient for all students and should contribute to developing and improving all the students' potentials toward a highest level, and that if the aim of schooling is to develop students who are harmonious and profitable to the people, then all communities should participate in assessment process. This also implies that in assessing students, it is necessary to look at this phenomena from multiple view points such as students, teachers, school management, communities and governments.

1.1. Need for Measurement and Assessment

Teaching and learning include a lot of instructional decisions to enhance and increase student learning, and quality of instruction is strongly connected to the structure of information on which these instructional decisions are made (Linn & Gronlund, 1995). Hence, the most important point is the determination of the way in which good, valid and reliable information about student learning can be provided. Traditionally, assessment process is focused on evaluating student accomplishment, however, contemporarily, it should be focused on increasing student learning and, the heart of assessment is a continuing flow in which the teacher in collaboration with students, uses information to guide the next steps in learning (Neill, 1997).

Scientific knowledge on student learning at science courses can be taken with multiple-measurement tools; for example, different kinds of testing methods are used to determine students' achievement levels and performance levels (Neill, 1997; Walker, 1999). Thus, science teachers have to know all these methods and be good evaluators of learners' progress. In addition, assessment has an important effect on learners' growth, achievement and self-esteem (Howe & Jones, 1998). In case of science teachers' using new and modern measurement and assessment methods throughout their teaching practices, it is believed that they can plan sequential learning activities for their students (Doran, Tamir & Chan, 1995). In this context, assessment methods such as project work, group work, higher levels of inquiries and using technological materials will increase the motivation of students to apply science knowledge into their own cognitive worlds (Doran, Tamir & Chan, 1995). This implies that a teacher should understand the particular student or group; that is to say, the teacher should assess students' actual strengths and learning needs, which require classroom-based performance assessment (Neill, 1997).

It is known that assessment takes place in three different forms throughout classroom teaching period: the beginning is assigned to finding out what students know and don't know; the instruction period is intended to help teachers make decision about what to do next, and the end of instruction is expected to provide teachers and students with an idea on what progress students have made and what they are capable of doing. Carin and Sund (1980: 274) explained this process as follows: science teachers determine what to teach - science content - at the beginning of a science course; then focus on how to teach- physical environment and intellectual levels of students, teaching methods, science materials; evaluate how much students have learned at the end of the course. If we look at these three basic phases in planning science courses, it seems that all of them, especially the last one, would require much more importance. If science teachers think a lot on the evaluation phase, they can be good at their teaching practices and profession.

Measurement and assessment concepts have different meanings; nevertheless, a group of teachers understand them as the same words. Here, two descriptions are given both for measurement and assessment concepts. Assessment is a term that includes a lot of procedures used to gain information related to student learning and formations of some value judgments about learning progress (Linn & Gronlund, 1995). In spite of the fact that measurement is limited to quantitative descriptions of students' behaviors, assessment can include both qualitative and quantitative characteristics of students' learning outcomes. If so, one tends to ask the question of what the aim of assessment is. Howe and Jones (1998) answered this question as follows: teacher-assessment is not intended to situate learners into some categories or to teach pupils to complete for some grades, but it aims to encourage and motivate students to achieve the cognitive, affective, psychomotor and social goals of a course. In the traditional approach, assessment is used to select students according to their cognitive abilities and to help them graduate from low-level schools to the higher ones. In this case, there appears an important problem, which is that too much emphasis is placed on grades and too little on helping students learn. These points, then, lead to a reduction in intrinsic motivation (Baker & Piburn, 1997).

1.2. Overview to the Measurement and Assessment Area in the World

At present, research on teaching and learning supports the movement toward new assessment approaches (NRC, 1996). Related literature includes a series of conceptual models that introduce the means and importance of student assessment from different perspectives (Linn & Gronlund, 1995). These are called placement, formative, diagnostic and summative assessments. Here we want to stress especially on formative assessment approach because it is more important than the others in learning contexts. For example, it can provide teachers with a considerable amount of information in order to improve each student's progress and develop teaching programs. Constructivist assessment, which entered the educational area and is supported by many educators, also includes formative assessment through classroom teaching process (Baker & Piburn, 1997). Educational assessment cannot be separated from teaching-learning progress; hence, it requires being an integral part of the teaching-learning process and provides some rich opportunities for student development and improvement.

Using effective measurement and assessment techniques ensures meaningful science learning for students and also affects the ways in which teachers teach and students learn. Mueller and his colleagues (2001) identifies that systematical assessment of student learning outcomes in science courses is supposed to change towards a realistic framework and classroom of the future is believed to evaluate realistic situations that require application of science concepts, principles and theories. In this way, students will not select the true answers from a set of multiple-choice questions, the best match from a set of concerning terms or decide whether a statement is true or false. Instead of this, learners will observe and explore the situation, discuss the observation notes with other students, make important judgments on a situation provided for them and contribute to construction of new knowledge. However, traditional assessment approaches with conventional tests are still dominant on the current educational area in the world (Baker & Piburn, 1997) and much of the information situated in conventional tests is not connected to the real lives or interests of students (Howe & Jones, 1998). Traditional assessment is just focused on assessing the attainments of cognitive objectives (Collette & Chiappetta, 1989) and mostly includes paper and pencil tests consisting of multiple choices, true false and completion items based on problems and a few easy or free response questions. These methods are limited in that they do not provide information about the attainment of many of the inseparable objectives of science teaching (Howe & Jones, 1998).

Another issue is whether teachers are well informed about knowledge on assessment methods and how this can be achieved. We suppose that teachers are not well informed about how to make effective formative evaluation and they are also faced with problems in making judgments on student learning. However, if teachers were allowed to gain sound formative skills and then develop appropriate assessment policies, the benefit for future years would be substantial (Black, 1986). Most of the teachers do not use assessment in right forms. Ulmer (2001) explains that just in case of decreasing the time and work demand on teachers, they can be much more efficient in formative assessment of critical thinking in problem-based learning. Nevertheless, it is still problematic whether teachers can develop themselves in formative assessments or not.

Teaching students in classrooms ensures teachers' professional development and leads them to learn, construct and find out new knowledge on how the best students can be assessed. This implies that not only should science teachers implement the current evaluation techniques effectively but also they should develop new evaluation techniques to improve and enhance science teaching in contextual areas. Problem-based instruction requires learners to rely on their own thinking to start to struggle with the faced problems. In order to enhance teachers' teaching for critical thinking, described also as reflective thinking (Farrell, 1999; Zeichner & Liston, 1996), it is necessary to improve the quality of student thinking. This is that we should include learners into assessment process and thus, students can be an integral part of it. In order to include them into the assessment process in appropriate ways, we can encourage them to become independent learners who can take responsibility for their own learning. In this context, we also need to make learners aware of the importance of critical thinking before teaching practices. The help of formative assessment activities could achieve this (Ulmer, 2001).

Individuals who are related to learners want to know how well their children do in schools (Amrein & Berliner, 2002). Science achievement outcomes of students should be regularly disseminated to parents, community and students in meaningful ways (ESRD, 2001). In this context, the systematical use of a wide range of assessment procedures provides an objective and comprehensive basis to report each student's learning progress (Linn & Gronlund, 1995). However, a lot of problems can be encountered in reporting how much learners have achieved. For example, in student choices and decision making for the curriculum, there are no useful tests in order to measure progress for all learners in a comprehensive classroom. Here, the term 'comprehensive classroom' means the classroom in which a range of students with different intellectual competencies is educated.

The other issue is the wide-scale and standard assessments. In the late 1990, the Third

International Math and Science Study (TIMSS) attempted to make comparisons between American students' math and science abilities and those of the students in other countries. Publishing the American students' unexpectedly low science test scores, educational professionals have argued over the validity of such wide-scale assessments (Frontczak, Kowalski & Brown, 2002). If so, what kinds of shifts are required in assessment approaches?

Educational assessment has two main purposes. The first is to help teachers design the instruction while the second is to contribute to learners in their progress. Baker and Piburn (1997) used the terms of traditional assessment and constructivist assessment. In traditional assessment, students' cognitive knowledge is determined by using a teacher-made or standard test such as multiple choices, true/false; fill in the blank and short-answer questions. In constructivist assessment, however, essays, practical examinations, papers, projects, questionnaires, inventories, checklist, portfolios, teacher observations, discussions and interviews are preferred for that purpose.

In contemporary approaches, it is stressed that using only limited means to determine pupil achievement is not enough. For this reason effective comprehensive assessment is advocated by some educators in assessing student learning in a good way (Mueller et al, 2001). They list those characteristics of effective comprehensive assessment as follows. It, he remarks:

- Is directly linked with course and daily outcomes and standards,
- Includes opportunities for reproduction of factual knowledge and the application and/or creative production of skills, knowledge, or concepts,
- Includes opportunities for self-assessment, goal setting, and personal/professional development through reflection,
- Includes a variety of tools which take into consideration the characteristics of the learners and the structure of the content,
- Includes performance-based tasks or projects that have a real-life context,
- Provides vehicles for specific feedback to the student about strengths and areas for improvement,
- Is congruent with methods of instruction.

It can be understood from these criteria that in effective comprehensive assessment, students are situated in the real life context and while they are struggling with the problems they encounter, they are observed. Besides this, it is explicit that students should be assessed individually rather than collaboratively.

1.3. Science Learning and Assessment

For the purpose of classroom instruction, assessment procedures which are used can be classified in terms of their functional roles: placement assessment in which a teacher determines prerequisite skills, course goals, and the best form of learning and formative assessment in which he/she determines the learning progress, provides feedback to reinforce learning and corrects learning errors. It is also important for a teacher to critically examine and systemically look at the phenomena that appear in the classroom; for example, a teacher's recording observations is believed to serve to alert him to some aspects of a student's learning or attitude that requires immediate attention. There is also diagnostic assessment in which he/she determines the causes - i.e. intellectual, physical, emotional and environmental - of persistent learning difficulties (Linn & Gronlund, 1995). Those examining a student's performance areas that need special attention seldom emerge. Accordingly, a teacher should decide to investigate the causes of this behavior more closely. In fact, especially while adjusting the learning to individual differences of all the learners in the science program, teachers make use of diagnostic data. Another category is summative assessment in which a teacher determines end of course achievement for assigning grades or certifying mastery of objectives (Linn & Gronlund, 1995). It is clear that the recorded phenomena concentrate on describing incidents of student performance over a period of time. However, the sequence of phenomena can serve as a record of the student's own development towards long-term goals such as lifelong learning, self-concept, cooperative learning, skill development, study skills, knowledge development, and interest (ESRD, 2001).

Now, it is time to discuss about how a science teacher can make quality assessment. We think, before all, a science teacher should determine the match between the most appropriate organizational methods and the type of student information to be gathered. Currently, based on the constructivist learning approach, individual assessment and portfolios seem to be the most useful methods to observe student achievement in science education (ESRD, 2001). Let’s look at the most important and differing features of these methods.

Individual assessment is mostly concerned with individual student progress and assessment activities constructed by the teacher are followed and completed individually by the students (ESRD, 2001). For example, teachers may wish to have students work individually on written assignments, presentations or performance assessment tasks in order to assess their individual progress at science lessons. Students' learning how to reflect themselves and how to evaluate their own works are quite important. Similarly, an important goal of school is for students to be able to learn without relying on teachers (Neill, 1997). In individual assessment, each student at each grade level is assessed according to his or her standing in attaining the objectives set out in curriculum documents and also self-referenced standards provide specific feedback to the individual on strengths and weaknesses. In addition, they are useful for motivating students and allow for a more relevant method of reporting progress for students with special needs and it can motivate students to accept a greater degree of responsibility for their learning progress (ESRD, 2001).

Portfolio is a collection of student-produced materials provided over an extended period of time that allows a teacher to evaluate student growth and overall learning progress during that period of time. It is an organizational structure that teachers may use to accumulate and organize student assessment information. Copies of assignments, contracts, assessments of presentations, assessments of the performance of skills and processes, quizzes, and tests are all examples of items, which may be included in portfolios (ESRD, 2001). In addition, a portfolio may also include samples of students' daily works. From a different viewpoint, portfolios are becoming popular among constructivists as one of the most appropriate forms of assessment. It also fosters the reflection and development of meta-cognitive strategies as students themselves evaluate and monitor their own progress (Champagne & Newell, 1992). The literature, which pertains to portfolios and portfolio assessment, highlights the learner's responsibility for selecting the work to be included in the portfolio and for the learning itself. Learners are not often given the opportunity to assess and judge their own learning. The portfolio, however, requires that learners take an active role in the evaluation of their own work and this helps to shift the emphasis from teaching to learning. The instruction and organization of the curriculum center around the portfolio and promote interactions between teacher and student around the collection of work (Klenowski, 1996). The process of putting together a portfolio results in a number of positive outcomes in addition to learning and development of positive attitudes towards science and scientific dispositions (Baker & Piburn, 1997)

Up to this point in the present article, we have explained in sum the trends and developments in the educational science assessment area in the world. In this context, we have discussed the issues of why science assessment is important, how science learning can be assessed in the best way. We have also considered some important features of two contemporary science assessment methods and so on. Next, we have explained how science learning is assessed in Turkish educational context. In addition, some significant problems faced during the science assessments of Turkish students have been considered.

2. Measurements and Assessment in the Turkish Context

In this section we have introduced the general structure of Turkish educational system and discussed the measurement and assessment approaches used in the Turkish educational system, examining the problems encountered in the assessment of Turkish students. And finally, a large problem source of assessment area, the university entrance examination (OSS), has been identified, and then we have focused on its validity and effectiveness as a standard and national exam.

2.1. Structure of the Turkish Educational System

Turkish educational system consists of four parts, which continue one after the other. Each of these is explained in details below:

2.1.1. Pre-school Education, an optional education system, aims at contributing to the physical, mental and emotional development of the children, helps them acquire good habits and prepares them for primary education. Pre-school education institutions include independent kindergartens, nursery classes in primary schools and preparation classes.

2.1.2. Primary Education provides children with basic knowledge and ensures their physical, mental and moral development in accordance with national objectives. It generally comprises the education of children in the 6-14 years age group. An eight years' primary education is compulsory for all Turkish citizens who have reached the age of six. This level of education is free of charge in public schools. There are also private schools under the state control.

2.1.3. Secondary Education encompasses two categories of educational institutions, namely general high schools and vocational-technical high schools (lycées) where a minimum of three years of schooling is implemented after primary education. The aims of secondary education are to provide students with a knowledge of general culture, to acquaint them with problems of individual and societal nature and to motivate them to find solutions; to instill in them the strength and knowledge to participate in the economic, social and cultural development of the country and to prepare them, in line with their interests and talents, for institutions of higher learning.

i. General high schools are educational institutions, which prepare students for institutions of higher learning. They implement a three-year program over and above primary education, and comprise students in the 15-17 years age group.

ii. Vocational-technical high schools provide specialized instruction with the aim of training qualified personnel. The organization and periods of instruction of these schools are different. Some of them have a four-year program in which case the schooling age is 15-18.

2.1.4. Higher Education: The purpose of higher education is to train manpower within a system of contemporary educational and training principles to meet the needs of the country. It provides high-level specialized education in various fields for students who have completed their secondary education. Universities comprising several units are established by the state and by law as public corporations having autonomy in teaching and research.

2.2 Review of the Literature on Measurement and Assessment Approaches in Turkey

2.2.1. Overview to the Turkish Educational Context: Some Problems and Causes

In the Turkish schools, traditional teaching methods are mostly used. Teachers of both primary and secondary school levels generally use presentation method while teaching their courses. Laboratories are not used as the primary learning centers of science and it is mostly stressed that the majority of science teachers also use traditional methods while teaching in classrooms and prefer demonstration and deduction methods while implementing their laboratory activities (Kaya, Çepni & Küçük, 2004; Pekmez, 2001).

Many researchers have investigated the issue of why science teachers are not willing to use laboratories in teaching their courses. It was found out that many factors influence the failure to use laboratory in schools; one factor is that activities done at the laboratories are not consistent with the questions types asked at the university entrance examination; the other factors can be listed as the lack of material in science laboratories, lack of laboratory experts in schools, inconvenient physical conditions of science laboratories, science program fully oriented with subject matter and students' negative attitudes towards laboratories. In addition, the other and maybe the most important factor is that science teachers do not have professional knowledge and skills for implementing laboratory activities properly (Çepni, Ayas & Akdeniz, 1995; Pekmez, 2001; Sahin, 2001). These practices all indicated that a teacher-centered approach is still dominant on the Turkish educational contexts at the beginning of the 21st century. This is expected to change with student-centered methods as stressed in the new science program called "*student-centered program*" for the primary school levels and have started to be implemented since 2000's.

Of the issues explained so far, the most important one is measurement and assessment of students. Turkish education system is based on behaviorist approach and target, and target behaviors are still dominant in the Turkish teaching programs even if in student-centered program. Universities and National Ministry of Education have started to reject traditional practices and stressed on the contemporary teaching methods, assessment methods; however, essays, short answer tests, multiple-choice tests are mostly used by teachers. In fact, national examinations such as secondary schools entrance examination (LGS) and university entrance examination requires and encourages this application. Teachers are expected to use assessment for two aims: to find out the extent to which students reach knowledge and skills level, and to determine how students gain these knowledge and skills or to determine why they fail to gain these (Turgut, 1992). For student-centered teaching methods in which students are playing an active role, that of using only written essay type and oral examinations[1] seems not to be enough for determining the achievement level of students. Instead of this, while teachers are assessing students' achievement, besides measurement results, they should see the factors such as how pupils join into classroom activities, their scientific attitudes (doing observation, doing research, scientific thinking), gained and exhibited ideas, taking responsibility, study in a group collaboratively sharing present

[1] Means that a student is asked a few questions about a special subject area and waited to response to them by explaining orally.

knowledge with others (MEB, 2000). In spite of the importance of these factors discussed in the primary science education program of Turkey, there are many research reports showing that when science teachers are to determine the cognitive levels of attitudes of their students, most of them use written exams, multiple-choice tests, oral exams, short-answer tests, match tests, true-false test and seldom project works (Özden, 1997; Turgut, 1994). For example in a study, done with thirty-nine physics teachers, it was found out that teachers use written examinations the most (Yigit, Saka & Akdeniz, 2000). This was regarded as tantamount to admitting that they aimed at the cognitive level of learning. Some researchers have discussed the instruction contexts and methods; and also measurement and assessment activities which are insufficient (Çelik, 2000; Özden, 1997; Turgut, 1994). However, in Turkey, concerning to this problem, for teachers not being developed during their pre-service education (Yigit, Saka & Akdeniz, 2000) and not being aware of assessment approaches seem to be the first factors. In addition, physical contexts of schools and classrooms, with crowded classroom and lack of materials, lower level students' cognitive abilities and incompetence of teachers are also important. We believe that provided that teachers act as researchers in their classrooms (Cohen & Manion, 1995; Mcniff, 1993; Küçük, 2002; Schön, 1983), they might perform the expected practices in their schools and classrooms. This is, maybe, the first and even the only solution to the current problems.

2.2.2. National Examination Issue: The Cases of OSS and LGS

As mentioned above, the most important problem in determining student learning in Turkey is believed by many educators to be the comprehensive examinations, such as OSS and LGS (Baykul, 1999; Çepni, 1993). Having graduated from secondary schools, the students who want to enter the university have to take the OSS and, at the same time, have to obtain the required grade for entering the demanded department. The OSS is a nation-wide exam held once a year in almost all the cities of the country and Lefkosa in the Republic of Northern Cyprus. Allocation of the students takes place according to the following criteria: scores obtained from the OSS, students' average high school scores and priority ranking of fields of study. Every year, students who want to enter any university find themselves in an examination race, which would take a long time. In this context, some students based on their parents' economical and social status take private courses and continue to a private education institute by paying a great deal of money. Some people resemble students getting ready for the OSS to racehorses that are trained just to reach the finish line as the first one. This situation makes assessment as competitive and comparative as explained also by Black (1986). However, learning is not a competition (Baker & Piburn, 1997).

In addition, studies in the Turkish educational contexts about measurement and assessment are based on the cognitive development and formal operational levels. When the questions asked at these examinations are compared with the questions asked at the primary or secondary schools, it is seen that they are quite different from each other. For example, in many studies that compare the questions, it was found out that while questions asked at school exams are the low level of Bloom taxonomy and Piagetian theory, questions at the comprehensive examinations are at high levels of Bloom taxonomy (Çepni, Ayvaci & Keles, 2001; Karamustafaoglu, Sevim, Karamustafaoglu & Çepni, 2003). In a study done by Kemhacioglu (2001), 1252 questions, which were asked at the high school physics course-1 by twenty physics teachers from different kinds of high schools, and 38 questions asked at the OSS between the years of 1999 and 2000 were investigated in detail according to Bloom taxonomy. It was found out that most of the questions asked by high school physics teachers are at the knowledge, comprehension and practice levels while the questions asked at the OSS are at the comprehension, practice and analysis levels of Bloom taxonomy. Here, it can be seen from the findings that there is not a close relation between teachers' examination questions and OSS questions. Classroom assessment forms of current Turkish science teachers are targeted more at the level of facts and comprehension than at higher levels. Some research findings in which questions asked by science teachers at the primary and secondary schools are analyzed according to Bloom's taxonomy supported this idea (Çepni, Ayvaci & Keles, 2001; Çepni, Bacanak, Özsevgeç & Gökdere, 2001; Morgil & Bayan, 1996) and also similar findings were observed in the related literature. For example, Zoller (1993), Zoller and Tsaparlis (1997) found out that chemistry teachers mostly asked very low cognitive levels of exam questions to assess their students at high school levels and their questions were usually at the first three levels of Bloom taxonomy. Despite the fact that many modern science educators stress application-level questions or science performances today, most of the questions from that era of testing and assessment focused solely on what was considered "scientific fact," with no connection to the lives of the students (Veronesi, 2000). That kind of assessment (quick recall of facts) has led the way to contemporary, divergent views on performance-based science assessment. While facts are important, teachers need to assess more complex cognitive processes. These include the abilities to apply knowledge in solving problems, to analyze complex situations to arrive at new solutions and to evaluate hypotheses and theories. Questions at the low levels of cognitive development only encourage students to memorize the facts and this hinders their intellectual development (Çepni & Azar, 1998). Students who continuously encounter with the lower level questions are directed towards the basic level of thinking. On the other hand, high-level questions are helpful for students to think more creatively and multi-dimensionally (Brualdi, 1998). Students understand that what is assessed is important and what is not assessed is unimportant (Baker & Piburn, 1997). We have explained above that currently some movements are seen in the Turkish educational area in which more learner-centered and, say, more constructivist learning approaches are being supported on the educational area. However, even in this situation, there appears an opposite case because standard and traditional forms of evaluation are still dominant in the current context. There is a growing recognition that standardized and other traditional forms of evaluation are inadequate for assessing constructivist teaching (Baker & Piburn, 1997).

In a study done by Çepni (2003), exam questions of university instructors who work at different science departments were analyzed according to the cognitive levels of Bloom Taxonomy. It was concluded from the comparisons of the findings that 81 % of the questions were at the first three levels and only 19 % of them were at the last three levels of cognitive domain and thus, cognitive levels of the questions were seen surprisingly rather low. This implies that the problems with traditional testing have not gone away yet. Those tests offer no solution to the educational needs of children. All the mentioned points stress that current assessment is thus at a crisis point: the present model is incapable of meeting real needs, and a new approach is not dominant in the Turkish context as also indicated for their educational context by Neill (1997).

Then we wonder if there is any other system, which looks like ours. We think it seems clear that Hong Kong and Turkey have an educational system that is quite similar to each other. Klenowski (1996) describes Hong Kong system as follows: The Hong Kong system is highly bureaucratic and centralized in nature and strongly influenced by its selective function; curriculum development policy follows a top-down and center-periphery approach; the education system is examination oriented. Some development and improvement movements are work like the Turkish context in Hong Kong; for example, educational research has been conducted over the past three years to analyze the processes of student self-evaluation and portfolio assessment and their impact on learning (Klenowski, 1996).

3. Future Trends for Measurements and Assessment of Science Education

Many professionals in field of education are becoming increasingly involved in the assessment activities and knowledge of the process. Here a model, which is designed to guide, and which can be implemented to assess most institutional and departmental assessment, plans or programs will be discussed in some details (Walker, 1999). This model proposes that an assessment of a plan or program, with the primary goal of enhanced student academic achievement and educational program improvement, should exemplify the following five principles (Walker, 1999):

- Mission and educational goals are reflected in the assessment process.
- The conceptual framework is effective.
- Institutional personnel are involved in the design of the assessment process.
- Data are collected through multiple measures.
- An assessment of assessment activities is established.

Present assessments beyond the classrooms should be changed for two basic purposes: to provide a richer and fairer means of assessment for these purposes and to remove the control the tests exert over classroom instruction and assessment (Neill, 1997). Classroom performance assessment requires thinking about the child and about the context in which the child is or is not successfully learning. There is overwhelming evidence that suggests a need to shift from relying solely on traditional summative assessments, such as written tests and term papers, to the ongoing use of alternative assessments, such as simulations, portfolios, case studies, performances, and projects (Wiggins & McTighe, 1998; Veronesi, 2000). In this context, non-traditional assessment methods but alternative assessments, such as the more student-centered instructional strategies, have been shown to increase student achievement (Fogarty, 1997). To use these alternative methods gives a clearer picture of what students actually understand rather than what they have just memorized. This is achieved when we as instructors and teachers, provide a variety of assessment methods and student choice. It is believed to allow students to demonstrate their knowledge using vehicles, which are responsive to their learning needs.

The last science education research reports of the American Association for the Advancement of Science's Project 2061: Science for All Americans (1989) and Benchmarks for Scientific Literacy (1993) have stressed the more authentic and performance-based science learning over simplified testing methods. Performance assessment allows teachers to see and hear what students know and can do with more clarity than is possible on a written exam. From the students' point of view, it is important because they find themselves engaged to a greater degree in the nature of scientific enterprise with this kind of assessment than the standardized tests. There has been an ever-increasing realization that the teachers are not getting students to think, and even if they are thinking, teachers have insufficient evidence to demonstrate it. It is also seen that more authentic science assessments such as those in which students can demonstrate what they know and can do, have begun to appear with increasing frequency in the science education literature (Veronesi, 2000).

New assessment standards are supposed to require that students be placed in situations where they are given parameters of a problem and invited to propose solutions (Baird, 1995). In this way, students are expected to do more than respond to pre-written prompts on tests. In addition, an effective test should focus on how well students can apply their knowledge of science and technology to raw materials so that value is added and the resulting product can be sold competitively in the real world (Baird, 1995). Doran, Tamir and Chan (1995) analyzed the changes in assessment of science classroom learning. They discussed that using group-administered tests and pencil-paper tests changed to tests, which require student selection or choice of answers performance including projects, reports, and portfolios. There is a need for a shift from testing every student with a simplistic exam to using a combination of classroom-based information and on-demand performance exams (Neill, 1997).

Thus, students will be invited to take a more active role in determining what they know and do in science learning environments. They are also expected to maintain portfolios that contain evidence of their skills, the quality of their writing, reflective thoughts on their personal strengths and weaknesses and goals. Some researchers discussed about student assessment already moving toward portfolio assessment for all students. While this type of assessment requires more time than machine scoring a multiple-choice test, it provides a richer record of students' abilities. It also places responsibility on students to examine their progress over time and to take pride in personal growth.

Stiggins (1991) thought that in the future, accurate science assessment would require establishing assessment goals, educating practitioners, as well as the lay public, about authentic assessment, and creating a clear vision of standards and expectations. In addition, Veronesi (2000) explains that two trends will continue: a) states can report on students' abilities in science, some forms of easy scoring, inexpensive surrogate for real science performance of knowledge and skills will most probably continue well into the first decade of the 21st century; b) in a counter-trend, teachers will continue to develop an appreciation of the merits of performance-based science assessment.

Up to now, all explanations have obviously stressed that future assessment would be situated on performance assessment, not on determination of achievement levels of learners. Thus, it is high time to think and begin to seek for solutions to the existing problems of our traditional education system so that we will adapt our traditional system to the modern times', namely contemporary, system. We should also be contemplative over the one, which achieves this mission at first and goes one or more steps further in the years to come. So we should make a choice between the two, either to go further more or remain stable.

References

American Association For the advancement of Science (1989). *Science for all Americans: A project 2061*. Reports on Literacy Goals in Science, Mathematics and Technology, Washington, DC.

American Association For the advancement of Science. (1993). *Benchmarks for science literacy*. New York, Oxford University Press.

Amrein, A. L. & Berliner, D. C. (2002, March 28). High-stakes testing, uncertainty, and student learning *Education Policy Analysis Archives, 10(18)*. Retrieved [Sept. 24, 2004] from http://epaa.asu.edu/epaa/v10n18/

Baird, B. (1995). The high school science classrooms of the future. *The High School Journal* (Special issue on the future secondary education. [Available from http://horizon.unc.edu/projects/HSJ/Baird.asp).

Baker, D. R. & Piburn, M. D. (1997). *Constructing science in middle and secondary school classrooms*. Boston: Allyn and Bacon.

Baykul, Y. (1999). İlkögretimde etkili ögretme ve ögrenme ögretmen el kitabi, ilkögretimde ölçme ve degerlendirme (Modul 3), [Available from http://www.meb.gov.tr].

Black, H. (1986). Assessment for learning, in assessing educational achievement. London: The Folmer Press, p.7-18 (Edited By Desmod L. Nuttall).

Brualdi, A. C. (1998). Classroom questions. *Practical Assessment Research & Evaluation, 6* (6), Eric Document reproduction no: ED 422407.

Carin, A. A. & Sund, R. B. (1980). *Teaching science through discovery*. Columbus, Ohio: Charles E. Merrill Publishing Co.

Çelik, D. (2000). *Okullarda ölçme ve degerlendirme nasil olmalidir?*. Milli Egitim Press, Istanbul, Turkey.

Çepni, S. & Azar, A. (1998). *The analysis of the physics questions asked at high schools exams*. III. National Science Education Symposium, Karadeniz Technical University, Trabzon, Turkey.

Çepni, S. (1993). *New science teachers' development in turkey: Implementation for the academy of new teachers' program*. Unpublished doctoral dissertation, University of Southampton, England.

Çepni, S. (2003). Analysis of university science instructors' examination questions according to the cognitive levels. Educational Sciences: *Theory & Practice, 3(1)*, Mayis / May 2003, 78-84.

Çepni, S., Ayas, A. & Akdeniz, A. R. (1994). Fen bilimlerinde laboratuarin yeri ve önemi (III): Ülkemizde laboratuar kullanimi ve bazi öneriler. *Çagdas Egitim Dergisi, 206*, 24-28.

Çepni, S., Ayvaci, H. S. & Keles, E. (2001, Eylül). *Okullarda ve lise giris sinavlarinda sorulan* *fen bilgisi sorularinin Bloom Taksonomisine göre karsilastirilmasi*. Science Education Symposium in Turkey At The Beginning Of New Thousand Year, University of Maltepe, Istanbul, Turkey.

Çepni, S., Bacanak, A., Özsevgeç, T. & Gökdere, M. (2001, Eylül). *LGS sorularinin ve fen bilgisi ögretmenlerinin sorduklari sorularin formal operasyon dönem özellikleriyle iliskilendirilmesi*. Science Education Symposium in Turkey At The Beginning Of New Thousand Year, University of Maltepe, Istanbul, Turkey.

Champagne, A. & Newell, S. (1992). Directions for research and development: Alternative methods for assessing scientific literacy. *Journal of Research In Science Teaching, 29*, 841-860.

Cohen, L. & Manion, L. (1995). *Research methods in education*, (Third Edit.), Routledge.

Collette, A. T. & Chiappetta, E. L. (1989). *Science instruction in the middle and secondary schools*. (Second Edit.), London: Merrill Publishing Company.

Doran, R. L., Tamir, P. & Chan, A. (1995). *Assessment in science.* Arlington, VA: National Science Teachers Association. pp. 17-18.

ESRD (2004). Student evaluation, *A teacher handbook*. Evaluation and Student Records Division of Saskatchewan Education. [Available from http://www.sasked.gov.sk.ca/docs/policy/studeval/introbac.html]

Farrell, T. S. C. (1999). Reflective teaching: A case study of a Korean English teacher. *Asian Journal of English Language Teaching 9*, 105-114.

Fogarty, R. (1997). *Problem-based learning and other curriculum models for the multiple intelligence classrooms*. Illinois: IRI/skylight Training and Publishing.

Frontczak, K., P., Kowalski, K. & Brown, R. D. (2002). Preschool teachers' use of assessments and curricula: A statewide examination, *Exceptional Children, 69(1)*, 109-123.

Howe, A. C. & Jones, L. (1998). *Engaging children in science*. Merrill, Prentice Hall, Inc., (Second Edition), Columbus, Ohio.

Karamustafaoglu, S., Sevim, S., Karamustafaoglu, O. & Çepni, S. (2003). Analysis of Turkish high-school chemistry-examination questions according to Bloom’s Taxonomy. *Chemistry Education: Research and Practice, 4 (1)*, 25-30.

Kaya, A., Çepni, S. & Küçük, M. (2004). Fizik ögretmenlerinin laboratuarlara yönelik hizmet içi ihtiyaçlari için bir program gelistirme çalismasi. *Gazi Üniversitesi Kastamonu Egitim Dergisi, 12(1)*, 41-56.

Kemhacioglu, T. (2001). *ÖSS sorularinin lise fizik-1 müfredati ile iliskisi*, Unpublished master dissertation, Karadeniz Technical University, Turkey.

Klenowski, V. (1996, September). *Connecting assessment and learning*. British Educational Research Association Annual Conference, Lancaster University.

Küçük, M. (2002). *Implementation of an in-service action research course program for science* *teachers: A case study*. Unpublished master dissertation, Karadeniz Technical University, Turkey.

Linn, R. L. & Gronlund, N. E. (1995). *Measurement and assessment in teaching*. (7th edition) Englewood Cliffs, NJ: Prentice-Hall.

Mcniff, J. (1995). *Teaching as learning an action research approach*. Routledge, London and New York.

MEB (2000). *Student-centered science teaching program of primary school*, Ankara, Turkey.

Morgil, F. İ. & Bayan, S. (1996). Success of secondary school students in solving physics questions of OSS and OYS examinations. *Hacettepe Journal of Education, 12*, 215-220.

Mueller, S. S., Waters, F. H., Smeaton, P. S. & Pinciotti, P. (2001). Making informed choices: A model for comprehensive classroom assessment. *Academic Exchange Quarterly, Spring 5(1)* 31

National Forum on Assessment, *Principles and Indicators for Student Assessment System*s, The National Center for Fair & Open Testing. [Available from http://www.fairtest.org/princind.htm]

National Research Council (1996). *National science education standards*. Washington, DC: National Academy Press.

Neill, D. M. (1997). Transforming student assessment. *Phi Delta Kappan, 79(1)*, 34-40.

Özden, V. (1997). *Ögrenme ve ögretme*. Pagem Press, (3. edit), Ankara, Turkey

Pekmez, E. S. (2001). *Ögretmenlerin fen egitiminde kullandiklari deneysel çalismalar ile ilgili görüslerinin incelenerek fen egitimi müfredat programlarindaki yerinin belirlenmesi*, Unpublished doctoral dissertation, University of Durham, UK.

Sahin, Y. (2001). *Türkiye'deki bazi üniversitelerin egitim fakültelerindeki temel fizik laboratuarlarinin kullanimi ve uygulanan yaklasimlarin degerlendirilmesi*, Unpublished master dissertation, Karadeniz Technical University, Turkey.

Schön, D. A. (1983). *The reflective practitioner*. London: Temple Smith.

Stiggins, R. J. (1991). Facing the challenges of a new era of educational assessment. *Applied Measurement in Education 4(4)*, 263-73.

Turgut, M. F. (1992). *Egitimde ölçme ve degerlendirme metotlari*. Saydam Press, (9. edit), Ankara, Turkey

Turgut, M. F. (1994, Haziran). *Fen ögretiminde ölçme ve degerlendirme*. 8. Türk Egitim Dernegi Bilimsel Toplantisi, Ankara,Turkey.

Ulmer, M. B. (2001). Self-grading for formative assessment in problem-based learning. *Academic Exchange Quarterly, Spring, 5(1)* 68.

Veronesi, P. (2000). Testing and assessment in science education: looking past the scoreboard.

The Clearing House, 74 (1), 27-30.

Walker, D. A. (1999). A model for assessing assessment activities. College Student Journal, Sept, 33(3) 439.

Wiggins, G. & McTighe, G. (1998). *Understanding by design*. Alexandria, VA: ASCD.

Yigit, N., Saka, A. Z. & Akdeniz, A. R. (2000). *Fizik derslerinde uygulanan ölçme-degerlendirme yaklasimlari,* III. National Science Education Symposium, Karadeniz Technical University, Trabzon, Turkey.

Zeichner, M. & Liston, P. (1996). *Reflective teaching: An introduction*. New Jersey: Lawrence Erlbaum.

Zoller, U. & Tsaparlis, G. (1997). Higher and lower-order cognitive skills: The case of chemistry. *Research in Science Education, 27(1)*, 117-130.

Zoller, U. (1993). Are lecturing and learning compatible? Maybe for LOCS: Unlikely for HOCS. *Journal of Chemical Education, 70 (3)*, 195-197.