MATHEMATICS AND SCIENCE ASSESSMENT IN THE TURKISH EDUCATIONAL SYSTEM
An Overview

Alipasa Ayas
Bilkent University

Emin Aydin
Marmara University

M. Sencer Corlu
Bilkent University

Changes in the nature of assessment in science and mathematics have led to a profound change in the conceptions of assessment. The change in our understanding of assessment reflected a move toward the greater integration of assessment and learning—away from assessment instruments whose links to learning were weak. The aim of the current paper was to review how these changes were reflected in the Turkish context, mainly in 2 parts: (1) governmental interventions, including changes in the middle grades teacher education, assessment system, and curriculum, and (2) the implementation of these changes in the middle grades classrooms. The review is based on three sources, including the program documents, textbooks, and external examinations. The analysis of these sources suggested that the program documents were the most successful in guiding mathematics and science teachers' assessment practices; however, several problems in the implementation of the prescribed ideas in the middle grades classrooms remained unresolved.

In the last 3 decades, the nature of assessment in science and mathematics has undergone a remarkable change (Kulm, 1994a, 1994b; Kulm & Malcom, 1994), reflecting a profound shift in the conceptions of assessment. There has been strong pressure to raise standards for all pupils. The emphasis on selecting the ablest students—through norm referencing—has been replaced by judging against a criterion, recognizing individual differences, and placing more emphasis on process than content and less emphasis on factual knowledge (Broadfoot, 1995).

There have been shifts in three major paradigms in educational psychology: psychometrics, educational measurement, and educational assessment (Gipps, 1994). The underlying philosophy of psychometrics is the

• **Emin Aydin**, Associate Professor, Ataturk Faculty of Education, Marmara University, Turkey. E-mail: eaydin@marmara.edu.tr

Middle Grades Research Journal, Volume 8(2), 2013, pp. 11–23 ISSN 1937-0814

conception of intelligence as a fixed and measurable trait. Norm-referenced tests are the basic source of information. The tenet of the educational measurement movement is its positioning in opposition to the idea of norm referencing. According to Wood (1986), the characteristics of the educational measurement movement are:

- criterion-referencing;
- competence rather than intelligence;
- uncontrollability of the measurement conditions;
- best rather than typical performance; and
- helping rather than sentencing the individual.

The core idea of educational assessment (third paradigm) uses assessment to support learning. Portfolios of accomplishments, situations which elicit problem-solving behavior, and appropriate scoring procedures fall into this category. A wider range of assessment techniques is needed to expand the range of cognitive skills.

The pressure for change that has led to this new paradigm comes from at least three sources. The first is a growing desire to broaden education, that is, to develop and consequently assess a much broader range of pupil abilities, which necessitates the use of a wide range of assessment techniques (National Council of Teachers of Mathematics [NCTM], 1995). The second is the desire to harness the full power of assessment and feedback in support of learning. The third arises from the belief that education should lead to a capacity for independent judgment and the ability to evaluate one's own performance, and that these abilities can only be developed through involvement in the assessment process.

The paradigm shift involves clarifying the purposes of assessment and establishing standards for judging the quality of assessment (NCTM, 1995; National Science Teachers Association [NSTA], 2013). The change reflects a move toward the greater integration of assessment and learning, and away from assessment instruments whose links to learning are weak. The paradigm shift in assessment continues to inspire hope that improvement in classroom assessment will lead to improvement in learning outcomes (Black & Wiliam, 1998).

THE TURKISH CONTEXT

An allusion to Rico's (1993) remarks on the interaction between culture and reform, Turkish assessment reforms have always been the products of the values of Turkish society. Similar to reform efforts in many other countries, Turkish curricular initiatives embodied primarily a top-down approach, but the society itself gave shape to the process. This process remained in effect until the next curricular intervention.

The Turkish education system experienced radical changes during the Republican period (1923–1940s). Beginning with the Unification of Education Act of 1924, there were significant changes in all aspects, including changes in its structure, program, and assessment schemes. Later, a slow process of innovation in the Turkish education system began nearly 2 decades ago with the aim of adjusting education to better fit the needs of the developing world. During this initiative, the curriculum was revised in consideration of innovations in technology, subject areas, educational sciences, and European Union standards—*EU acquis*.

The 1998 program document (the old curriculum) was a slightly revised form of the 1983 version (Ministry of National Education [MoNE], 1983), which came right after the extension of compulsory public education from 5 to 8 years (MoNE, 1998). The 1998 document contained statements or *learning objectives*, most of which came from the first two levels of Bloom's taxonomy—knowledge and understanding. Although the 1998 document did not specifically address the assessment issue (Alkan, 2002; Albayrak & Aydin, 2002), curricula renovation was its top agenda.

Specialized commissions were established and a needs analysis was conducted for each school subject—from elementary through secondary levels—in line with the opinions obtained from nongovernmental organizations, universities, inspectors, administrators, and teachers. The coverage of school subjects was something that had never happened before. All school curricula were included in this overhaul—from social sciences to physical sciences, mathematics to arts, and physical education.

The last wave of these aforementioned changes in the Turkish context emerged a few years before the new millennium; first, as an awareness in the academia, followed by the 2005 curriculum reforms (Aydin & Delice, 2010). The role of the government in changing the teacher education system, assessment system, and curriculum, and in implementing the changes in the classroom context was noteworthy. Because there have been no tradition of internally assessing teachers' performances, government influence focused on changes in the curriculum and preservice and in-service teacher training programs. In evaluating the influence of the government on reforms, which began with the National Education Development Project (1992-1998) and ended with the 2005 curriculum reform, the authors of this paper examined research on three primary data sources:

- Program documents: to understand the intentions of the curriculum developers;
- Textbooks: to understand how well the intentions were put to immediate use by the teacher in the classroom; and
- External examinations: given that teachers *teach to the test* (Madaus, 1988), the authors believe that the question types can have a defining role in the curriculum and use of assessment tools and approaches.

2005 Curriculum Reforms

The changes in the 2005 document (new curriculum) replaced the concept of *learning objectives* by the notion of *attainment*, which was considered a move toward the new paradigm. The new curriculum document included not only the names of topics and general objectives that were in the traditional curriculum, but also specific learning outcomes phrased as *gained skills and abilities*. The new science and mathematics curricula contained teacher guides, student workbooks, and suggested ideas for performance-based assessments, as well. The aim of the reformed curriculum was to engage students in deeper learning.

The new science and mathematics curricula considered what students should learn in the classroom. There were fewer topics than in the previous curricula (MoNE, 2005, 2006, 2009). Another important difference between the old and new science and mathematics curricula was the phrasing of learning outcomes. In the old science curriculum (for elementary, Grades 4-8), a behavioral approach had been used. The goals, aims, and behavioral objectives for teaching science for each topic had been stated in the documents. In the new curriculum, these were phrased as *students' gains* (MoNE, 2009), reflecting a paradigm shift from educational measurement (Gipps, 1994) to educational assessment toward a more student-centered and constructivist view.

Assessment in the 2005 Program Document

In this section, the analysis focused on the *intended curriculum* (Cuban, 1992) as it was reflected in the curriculum documents for mathematics (MoNE, 2009), science (MoNE, 2005, 2006), and the MoNE approved textbooks. Given that the 2005 program document contained a separate section on assessment, the analyses presented in this paper is primarily based on the information provided in this assessment section.

In the 2005 program document, there were implicit references to the constructivist learning paradigm (MoNE, 2005). Hence, the expectancy was to alter assessment practices in line with this paradigm. In the document, a

clear distinction existed between the informal and formal uses of assessment tools. There was also awareness that the main purpose of classroom assessment was formative (NCTM, 1995; National Science Teachers Association, 2013)—that is, (1) to help the students in their individual learning processes and (2) to inform teachers about the influence of their own classroom actions (MoNE, 2005, 2006, 2009).

The document addressed the issues regarding changes in internal and external assessment. As to internal assessment, the document seemed to give equal importance to traditional (e.g., multiple choice, matching type, and short-answer tests) and nontraditional authentic tasks (e.g., journals, portfolios, control lists, and interview forms) (Capraro & Corlu, 2013). The document also aimed to increase the repertoire of tools that teachers could use cumulatively. Portfolios, projects, and performance tasks could be used for assessment, apart from the traditional paper-and-pencil tests (MoNE, 2005, 2006, 2009). However, the use of traditional assessment tools, such as multiple choice test items, were not completely disregarded. This seemed to be a realistic stance, given the prevailing testing culture developed over the years.

One of the assumptions of the current paradigm was that assessing performance is not an exact science and it involves the complex interaction between the student, task, and context (Gipps, 1994). The 2005 document exhibited an awareness of this *complex reality of the individual* by listing some of the student characteristics that could be assessed. For example, some student characteristics that the mathematics curriculum aimed to enhance were:

- the ability to use mathematics in daily life;
- mathematical reasoning skills;
- problem-solving skills;
- attitudes toward mathematics;
- self-regulation skills;
- social skills;
- aesthetic views; and
- mathematics communication (MoNE, 2009).

Similar goals existed in the science program document:

- to enable students to develop themselves in a way that they are scientifically literate;
- to equip the students to learn and understand the natural world, and enjoy living with its richness and enthusiasm;
- to encourage student curiosity about scientific and technologic developments and events;
- to associate and understand the relationship between science, technology, society, and the environment;
- to enable students to structure new knowledge by reading, searching, and discussing;
- to help students develop knowledge, curiosity, attitudes, and experience about science and science-related professions or jobs;
- to help students learn how to learn and follow the changing nature of knowledge and jobs for them to be able to update their knowledge and skills in their profession (MoNE, 2005, 2006).

The emphasis in the document on, *assessment as an integral part of the process of learning,* indicated a move from a formal to an informal perspective on assessment (Rowntree, 1987). *The encouragement of the use of a variety of tools* in every opportunity that could be found was in line with the assessment principles stated in NCTM (1995), NSTA (2013), and the National Assessment Governing Board (2008); that is, the *use of multiple sources of assessment* to reveal the complex reality of the student as much as possible. Tables 1 and 2 show two different examples from mathematics and science subjects.

The second aspect of governmental intervention was related to the way that document's principles were reflected in the textbooks. For example, when they examined Turkish textbooks vis-à-vis the textbooks used in the United States and Singapore, Ozer and Sezer (2012) found that Turkish textbooks contained significantly fewer questions, although the

TABLE 1
Attainment Targets Versus the Mental Processes Involved
From the Eighth-Grade Algebra Learning Area

Attainment Targets	*Mental Process (Based on Bloom's Taxonomy)*
Define rational numbers.	Comprehension
Do addition and subtraction with rational numbers.	Application
Do multiplication with rational numbers.	Application
Do division with rational numbers.	Application
Estimate the result of an operation using a strategy.	Analysis
Solve and construct word problems involving rational numbers.	Analysis

TABLE 2
Attainment Targets and the Mental Process Examples From 12th-Grade Physical Science

	Attainment Targets	*Mental Process (Based on Bloom's Taxonomy)*
1.	Give an example of an element that has an isotope.	Knowledge
2.	Predict some common chemical reactions, given a choice of reactants (e.g., metals and nonmetals, acids and bases).	Comprehension
3.	Explain the difference between ionic and covalent bonding.	Application
4.	Design an investigation to determine the effect of surface area on evaporation rate.	Evaluation
5.	Design an investigation to measure the temperature of water when it changes state (phase diagrams).	Analysis

questions were more complex and involved higher mental processes. In order to illustrate the differences in question types before and after the governmental interventions, two question types from the same grade and learning unit (sets) taken from a new and old textbook are presented. See Figure 1.

It is evident in Figure 1 that the questions from the old textbook were mainly of the noncontextual drill and practice type, which could only assess the skill of applying a routine procedure. The questions in the new textbook were contextual problems that necessitated knowledge transfer and the application of existing knowledge to new contexts.

Changes in the assessment system. Despite all the positive aspects, the changes concerning the assessment system have been about the external *high stakes* testing; *increasing the quality of the learning outcomes* has rarely been the issue. There were no concerns about the format of test items (multiple choice vs. open response), or the level of content in regard to Bloom's or any other taxonomy. The changes were primarily limited to the number of examinations that a student had to pass for high school or university placement, under which the basic argument was the *difficulties arising from determining a student's future in a 3-hour examination*.

The dominant teacher perception before the curriculum reform was the separation between the processes of teaching and assessment. This could partly be attributed to the fact that Turkish teachers were overburdened from having to cope simultaneously with the formative and summative functions of assessment. Taking into consideration the dominant character of

① Filiz Öğretmen 6 A sınıfındaki öğrenci velilerinin meslekleri ile ilgili bilgi edinmek için bir anket yapar. Anket sonuçları aşağıdaki tabloda verilmiştir:

Tablo:Velilerinin Meslekleri

Öğretmen	||||
Doktor	||
Avukat	|
İşçi	𝍸 ||
Manav	||
Terzi	
Polis	|||
Serbest	𝍸 𝍸
Diğer	||

a)Tablodaki meslek gruplarını birer küme olarak düşünüp bu kümelerin eleman sayılarını sembolle gösteriniz.
b)Velisi terzi olanların kümesini nasıl gösterebilirsiniz?
c)Okulun tamamında bu şekilde bir çalışma yapılsaydı evrensel küme ne olurdu?

Siz de bu ankete katılmış olsaydınız hangi kümenin eleman sayısında değişiklik olurdu?

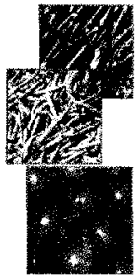
Birleşim ve Kesişim İşlemi

Aysel Hanım misafirleri için yemek hazırlayacaktır. Mutfağa gittiğinde elindeki malzemelerle sadece patlıcan veya fasulye yemeği pişirebileceğini görür. Çünkü evdeki domates, soğan ve biber bir yemek için yeterlidir.

Evde bulunan malzemeler
patlıcan
fasulye
domates
soğan
biber
sarımsak
dolmalık biber
salatalık
yağ

Patlıcan yemeği için gerekli malzemeler
patlıcan
domates
soğan
biber
sarımsak
yağ

Fasulye yemeği için gerekli malzemeler
fasulye
domates
soğan
biber
yağ



a) Listeye göre her iki yemekte kullanılabilecek ortak malzemeler hangileridir?
b) Her iki yemek için gerekli olan tüm malzemeler hangileridir?
c) Patlıcan yemeğinde kullanılıp fasulye yemeğinde kullanılmayan malzemeler nelerdir?
ç) Aysel Hanım'ın evindeki malzemelerden patlıcan veya fasulye yemeğinde kullanılmayanlar hangileridir?

EXERCISES 2.5

1. List all the subsets of each of the following

   a) {p} b) {p,q} c) {p,q,r}

2. Use 2^n, find the number of subsets

   a) {a} b) {a,b} c) {a,b,c}

   d) {a,b,c,d,f} e) {a,b,c,d,e,g}

3. Use 2^n-1, in problem two, find the number of proper subsets.

4. A = {3,6,9,12}
   C = {4,8,12}
   E = {2,4,6,8}

   State whether the following are true (T) or false (F)

   a) $A \subseteq B$ b) $D \subset C$ c) $C \not\subseteq E$

   d) $C \subseteq E$ e) $D \subseteq A$ f) $D \not\subseteq B$

   g) $C \subseteq A$ h) $A \subseteq E$ i) $E \subseteq B$

   j) $D \subseteq C$ k) $\{6,9\} \not\subseteq A$ l) $\{4,8\} \subset B$

   m) $\{6,12\} \not\subseteq E$ n) $\{2,4\} \subseteq C$ o) $\{5,7,9,11\} \subset B$

   p) $\{6\} \subseteq D$ q) $\{\ \} \not\subseteq A$ s) $\{4,8,12\} \subset C$

5. List all the subsets of the following

   a) T = {a, {b} }

   b) K = { {1}, {Ø}, 3}

6. P,Q and R are three non-empty sets with the properties $P \subset Q$,

   a) $Q \subset R$

   b) $P \subset R$

   Draw a Venn diagram to show the relationship between the sets.

FIGURE 1

Question Examples From the Set Learning Unit Taken From a New (Durmus et al., 2008) and Old (Baglan et al., 1986) Sixth-Grade Textbook

the external assessments, exam preparations seemed to steal most of the time allotted for actual teaching. Pressure from parents and school administrators were compelling teachers to act as exam coaches rather than educators. The Turkish external assessment system has been highly selective and placement-oriented, so that its influence has been more decisive on classroom assessment practices than on any other factors, including curriculum changes and teacher training programs (Ayas, 2012).

At this point, the attention is on the external examinations; specifically, the external assessment system used in the middle grades. For selection and placement to higher level institutions after middle school, there was only one examination, the Secondary School Student Selection and Placement Examination (OKS), which was given at the end of the eighth grade. This was revamped in 2008 to make the summative assessment more compatible with classroom teaching. The new system was called the Examination for Level Determination (SBS). The name suggested a more *formative* approach to assessment compared to the former approach, which had been more focused on summative assessment. In the new system, summative assessments were conducted at the end of the sixth, seventh, and eighth grades. Each assessment contained 100 questions, equally distributed across the subject areas of Turkish grammar, mathematics, science, and social science.

OKS was successful in fulfilling the needs pertaining to *measuring*, but not successful enough to *assess* all aspects of students' academic achievement (MoNE, 2008). The negative psychological effects resulting from making placement decisions based on a single *terminal* examination could have been minimized by increasing the number of examinations (MoNE, 2008). Therefore, passing from OKS to SBS signaled a move from a measurement paradigm to an assessment paradigm (Gipps, 1994). The intention might have been justified, but it was hard to tell whether the SBS system has achieved these goals. In fact, there appeared official statements that the summative functions loaded into SBS would completely vanish and the examination would only carry the purpose of providing feedback to the system (Radikal, 2012).

With the broader aim of determining whether or not the external examinations conducted in the middle grades successfully assessed higher order thinking skills, Ugurel, Morali, and Kesgin (2012) analyzed the test items of OKS and SBS with respect to the MATH (Mathematical Assessment Task Hierarchy) taxonomy of Smith et al. (1996), a revised version of Bloom's Taxonomy. See Table 3 for its groups and categories.

According to Ugurel et al. (2012), most of the questions were in categories A and B, indicating that the examinations predominantly assessed the skills of routine procedures and knowledge transfer (see Figure 2). The questions only required the application of memorized and practiced algorithms (A3), and the transfer of knowledge from one form of representation to another (e.g., algebraic to graphical). The researchers also compared the questions of both examinations with those from the Trends in International Mathematics and Science Study (TIMSS). They found that with respect to MATH levels, the SBS questions were closer to those of the TIMMS than the OKS questions were.

Changes in the teacher education programs. In order to overcome the difficulties related to the internalization of the contemporary assessment paradigm and effective use of nontraditional tools, preservice and in-service teacher training programs must be given extra attention. The World Bank-funded National Education Development Project (1992-1998) (The British Council, 2005) can be considered a turning point for implementing international standards in preservice teacher training programs. The objective of this project was to improve preservice teacher education at the primary and secondary levels by (1) raising the standards in teacher education and establishing an accreditation system, (2) providing long- and short-term fellowships, (3) upgrad-

TABLE 3
Groups and Categories of the MATH Taxonomy

	MATH Taxonomy
A1	Knowledge and knowledge systems
A2	Comprehension
A3	Use of routine procedures
B1	Transfer of knowledge
B2	Application to new contexts
C1	Verification and interpretation
C2	Deduction, estimation, and comparisons
C3	Evaluation

	SBS-6	N	SBS-7	N	SBS-8	N	OKS	N
A.1	--	0	--	0	--	0	--	0
A.2	5,13	2	2,16	2	2,4,10,13	4	1	1
A.3	1,10,12,15	4	1,3,5,7,8,10,12,14	**8**	1,6,9,11,14	**5**	2,6,7,9,15,17,24	7
B.1	2,3,6,7,11,16	**6**	6,9,13,17	4	3,7,12,15,16	**5**	4,5,18,21,22, 23,25	7
B.2	4,9,14	3	4,11,18	3	5,8,17,19	4	3,8,10,11,12, 13,14,16,19	**9**
C.1	--	0	--	0	18,20	2	20	1
C.2	8	1	15	1	--	0	--	0
C.3	--	0	--	0	--	0	--	0

FIGURE 2
Distribution of the Questions in the SBS and OKS Examinations With Respect to MATH (Ugurel et al., 2012).

ing the facilities of education faculty, and (4) developing student-teacher experience in schools (Council of Higher Education, 1997, 1999).

The restructuring of in-service training programs in most areas of primary and secondary education was completed by 1998. The young researchers who had been sent to the United Kingdom and the United States returned to their universities in the 1999–2002 interval, after having received graduate degrees in all areas of education. From 1996 onwards, teachers of mathematics and science (from elementary through secondary schools) have been educated with a standard program, in which the courses fostering content knowledge (mathematics or science), pedagogical content knowledge (mathematics or science education), and pedagogy knowledge (general education) were evenly distributed across the curriculum.

The reflection of this paradigm shift in teachers' conceptions about assessment can be a slow process (Black, 1993). Since the majority of teachers working today received their training before this restructuring occurred, their training did not address *pedagogical content knowledge* well-enough (Shulman, 1987). In-service training programs could be beneficial in addressing this deficiency that naturally occurs during a transmission period, but they have shown limited effectiveness except in a few attempts (Bingolbali, Akkoc, Ozmantar, & Demir, 2011).

Assessment in the Classroom

For an in-depth understanding of what the inside of the classroom was like before the 2005 program, a review of literature can be illuminating. However, there has been little data from inside the mathematics and science classrooms. In evaluating the influence of governmental interventions on teachers' assessment practices, the main obstacle was the dearth of literature from inside the classroom. Upon reviewing the papers presented in the 10th National Congress on Science and Mathematics Education, authors of the current review found that participants in most studies in mathematics education were teacher candidates; only 22 out of 133 papers were about practicing teachers, of which only four contained assessment-related data.

The earliest paper on the assessment practices in the classroom was published in 1996, in which the first signs of awareness of the role of assessment in learning were apparent (Alkan, Sezer, Ozcelik, & Koroglu, 1996). Researchers discussed general issues regarding the role of assessment in mathematics teaching and reported findings from a survey of a sample of 1,150 students and 175 teachers. The findings suggested that (1) there was a lack of classroom dialogue between teachers and students (which can be taken as an indication of the scarcity of informal teacher assessments), (2) external examinations were the dominant factor in determining classroom discourse, and (3) (as an extension of this) multiple-choice tests were the primary assessment tool. The findings of Albayrak and Aydin (2002) were consistent: assessment practices of middle school mathematics teachers were mostly summative and teachers did not show any effort to use the assessment results diagnostically. Apart from their regular summative examinations, the only assessment tools they used were multiple-choice tests produced by the *dershanes* (private tutoring institutions that prepare students for external examinations) (Oztelli, Corlu, Corlu, & Capraro, 2011).

In another study, Alkan (2002) reviewed the written records of ten mathematics departments in the district of Izmir, Turkey. The researcher found that of all the *target behaviors* that were chosen, only 22% were deemed measurable, all of which measured the skill of doing operations. Turnuklu (1993) identified three behavioral patterns in the assessment practices of Turkish mathematics teachers that related to their lack of skills in using assessment tools for recording information and using the information to monitor students' learning:

- *Paying more attention to whether students' homework is done rather than how students answer questions:* Two potential reasons were proposed for this behavior were not knowing how to assess them and not being able to ask the appropriate questions (e.g., drill and practice equations, which inhibit the observation of the process).
- *Difficulty in assessing term projects:* Teachers gave projects that were too difficult for the students' age group and were not sure whether the work was done by the students themselves; indicating a mistrust in their own grading.
- *Unsystematic recording of assessment results:* Thus, teachers relied more on memory and the grades given to students' written examination papers.

In a study conducted by Karacaoglu, Durmus, and Bal (2012), teachers seemed to use performance tasks formatively and cumulatively, but experienced time-management problems in the classroom. Kogce and Baki (2002) studied teachers' skills in using feedback in the classroom and found that teachers generally use verbal feedback, as it is allegedly easier to manage and takes less time. The findings also suggested that teachers were not eager to record the information obtained from their informal assessments. Bal (2012) reported similar findings for portfolio assessment.

Teacher Candidates

The review of the literature on the effectiveness of teacher training programs did not yield favorable results for the development of fundamental assessment skills, either. Findings from Akkoc's (2012) study indicated that the formative assessment practices that mathematics teacher candidates most frequently use were asking questions during teaching or giving homework. Mathematics teacher candidates seem to have difficulty solving questions that require the application of their knowledge to new situations (Kesgin & Moralli, 2012). The students' main method of preparation for the exam in the abstract algebra course was studying questions from past exams and trying to solve questions that were similar to those given in the course. This, we believe, might have unfavorable consequences in their future teaching practice. Kula and Bukova-Guzel (2012) identified two patterns in the questioning techniques of mathematics teacher candidates: asking questions that only require procedural knowledge and having to answer the questions they asked.

CONCLUSIONS

In examining teachers' assessment practices before and after the changes brought by the two important government-led interventions, several points emerge. There is a lack of knowledge and skill, especially concerning assessment in mathematics and science; therefore, the assessment skills that teachers have been mostly based on their experiences gained during their own preuniversity education and those learned from their colleagues during their professional practice. As a direct result of this, teachers are mostly unaware of the role of classroom assessment in learning mathematics and science. Teachers were also caught between the requirements of the intended curriculum and the de facto curriculum defined by the external examinations (Madaus, 1988, p. 93).

The ideas presented also seem to be sound in comparison with the curricula in Western developed countries. Overall, the program documents can be considered successful in guiding teachers' assessment practices. However, the implementation of the prescribed ideas in the Turkish schools and inside classrooms is problematic (Bal, 2012; Kula & Bukova-Guzel, 2002; Karacaoglu et al., 2012). The importance of assessment is nicely stated in the *Book Differentiated Assessment Strategies*, in the foreword by Marti Richardson, with a quotation from Michael Fullan's (1991) idea that "assessment has to drive the educational change agenda around learning and student achievement."

The policy-level changes that occurred in Turkey's science and mathematics curricula with the renewed assessment strategy did not seem to be translated into meaningful change in the classroom. However, for a number of reasons (including budgetary limitations), most teachers did not receive training on the innovations in curricula and assessment strategies. Also, the yearly nationwide external examination by the Student Selection and Placement Center to enroll students at the universities was not changed to reflect the new assessment strategies. The exam usually assesses the students' procedural knowledge with multiple-choice questions. There are high-level questions in the exam, but because there are courses that prepare students for it, they are able to do many exercises similar to the questions. Thus, it just becomes a matter of memorizing the procedure for solving such types of questions. Therefore, the nationwide exam does not drive the changes made in assessment in the implementation of the science curricula in schools.

Teachers need in-service training to renew their knowledge, skills, and ideas about assessment. The changes in preservice teacher training programs that paralleled those in the systems of most of the developed countries looked sound on paper. However, it appears that these programs have not been very effective so far (Bingolbali et al., 2011). Even the

quality of in-service training to teach the skills required for the new curricula is questionable (Elmas & Geban, 2010; Gunes et al., 2010; Senel-Coruhlu, Er Nas, & Cepni, 2009).

REFERENCES

Akkoc, H. (2012, June). *Matematik ogretmen adaylarinin bilgisayar destekli olçme degerlendirme pratikleri* [Computer-assisted assessment practices of mathematics student teachers]. Paper presented in X. Ulusal Fen Bilimleri ve Matematik Egitimi Kongresi, Nigde, Turkey.

Alkan, H. (2002, September). *Matematik egitiminde belirlenen hedef davranislar ile kullanilan olcme araçlarinin iliskisi* [The relationship between predefined target behaviors and the assessment tools used in mathematics education]. Paper presented in V. Ulusal Fen Bilimleri ve Matematik Egitimi Kongresi, Ankara, Turkey.

Alkan, H., Sezer, M., Ozcelik A. Z., & Koroglu, H. (1996, September). *Matematik egitiminde olcme ve degerlendirmenin etkisi* [Effect of assessment on mathematics education]. Paper presented in II. Ulusal Egitim Sempozyumu, Istanbul, Turkey.

Albayrak, M., & Aydin, Y. (2002, September). *1983'den 2002'ye ilkogretim matematik dersi programi* [Primary school mathematics teaching program from 1983 to 2002]. Paper presented in V. Ulusal Fen Bilimleri ve Matematik Egitimi Kongresi, Ankara, Turkey.

Ayas, A. (2012). An examination of Turkish science curricula from a historical perspective with an emphasis on learning outcomes. In S. Bernholt, K. Neumann, & P. Nentwig (Eds.), *Making it tangible—Learning outcomes in science education* (pp. 399-424). Munster, Germany: Waxmann.

Aydin, E., & Delice, A. (2010). Olcme-degerlendirmeye kavram yanilgilari perspektifinden bir bakis [An overview of assessment from the perspective of conceptual mistakes]. In M. F. Ozmantar, E. Bingolbali, & H. Akkoc (Eds.), *Matematiksel Kavram Yanilgilari ve* Çozum *Onerileri* (2nd ed., pp. 393-436), Istanbul, Turkey: Pegem Akademi.

Baglan, H., et al. (1986). *Modern mathematics one*. Istanbul, Turkey: Haset (Hachette) Kitabevi.

Bal, A. P. (2012, June). *Matematik ogretmenlerinin portfolyo degerlendirme surecine iliskin gorusleri* [The views of mathematics student teachers on the portfolio assessment process]. Paper presented in X. Ulusal Fen Bilimleri ve Matematik Egitimi Kongresi, Nigde, Turkey.

Bingolbali, E., Akkoc, H., Ozmantar, M. F., & Demir, S. (2011). Pre-service and in-service teachers' views of the sources of students' mathematical difficulties. *International Electronic Journal of Mathematics Education, 6*(1), 40-59.

Black, P. (1993). Formative and summative assessment by teachers. *Studies in Science Education*, *21,* 49-97.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy and Practice*, *5*(1), 7-75.

The British Council. (2005). *National Education Development Project, Turkey*. Retrieved from http://www.britishcouncil.org/development-education-experience-turkey-national-education.pdf

Broadfoot, P. (1995). Performance assessment in perspective: International trends and current English experience. In H. Torrance (Ed.), *Evaluating authentic assessment: Problems and possibilities in new approaches to assessment* (pp. 9-43). Buckingham, England: Open University Press.

Capraro, R. M., & Corlu, M. S. (2013). Changing views on assessment for STEM project-based learning. In R. M. Capraro, M. M. Capraro, & J. Morgan (Eds.). *STEM project-based learning: An integrated science, technology, engineering, and mathematics (STEM) approach* (2nd ed., pp. 109-118). Rotterdam, The Netherlands: Sense.

Council of Higher Education. (1997). *Fakulte-okul işbirligi* [Faculty-school partnership]. Ankara, Turkey: World Bank.

Council of Higher Education. (1999). *Turkiye'de ogretmen egitiminde standartlar ve akreditasyon* [Standards and accreditation teacher education in Turkey], Ankara, Turkey: World Bank.

Cuban, L. (1992). Curriculum stability and change. In J. Philip (Ed.), *Handbook of research on curriculum* (pp. 216-247). New York, NY: Macmillan.

Durmus, S. et al. (2008). *Matematik Ders Kitabi* [Mathematics textbook]. Ankara, Turkey: Milli Egitim Bakanligi.

Elmas, R., & Geban, O. (2010, June). *High school chemistry teachers' views related to the new chemistry curriculum*. Paper presented at World Council of Comparative Education Societies Conference, Istanbul, Turkey.

Fullan, M. G. (1991). *The new meaning of educational change*. London, England: Cassell.

Gipps, C. V. (1994). *Beyond testing: Towards a theory of educational assessment.* London, England: The Falmer Press.

Gunes, B., Ates, S., Eryilmaz, A., Kanli, U., Serin, G., Arslan, A., & Gulyurdu, T. (2010, September). *Yenilenen fizik dersi ogretim programinin uygulanma sureci ve yaşanan sikiintilar uzerine bir arastirma* [An investigation on the practice of the renewed physics teaching program]. Paper presented at the IX. Ulusal Fen Bilimleri ve Matematik Egitimi Kongresi, Izmir, Turkey.

Karacaoglu, A., Durmuş, D., & Bal, A. P. (2012, June). *Matematik ogretmenlerinin performans gorevi hazirlama, degerlendirme ve uygulama sureci hakkindaki gorusleri* [The views of mathematics teachers on the preparation, evaluation, and implementation of performance tasks]. Paper presented in X. Ulusal Fen Bilimleri ve Matematik Egitimi Kongresi, Nigde, Turkey.

Kesgin, S., & Moralii, S. (2012, June). *Matematik ogretmen adaylarinin MATH taksonomi cercevesinde hazirlanan sorulara iliskin gorusleri ve soru ornekleri* [The views of Turkish mathematics teachers on the questions prepared using MATH taxonomy and question examples]. Paper presented in X. Ulusal Fen Bilimleri ve Matematik Egitimi Kongresi, Nigde, Turkey.

Kogce, D., & Baki, A. (2012, June). *Ilkogretim matematik ogretmenlerinin geribildirimin veriliş tarzi ve zamanlamasi ile ilgili inanislari* [The views of primary school mathematics teachers on the styles and timings of feedback given to students]. Paper presented in X. Ulusal Fen Bilimleri ve Matematik Egitimi Kongresi, Nigde, Turkey.

Kulm, G. (Ed.). (1994a). *Assessing higher order thinking in mathematics.* Hillsdale, NJ: Erlbaum (Original work published 1990)

Kulm, G. (1994b). *Mathematics assessment: What works in the classroom*. San Francisco, CA: Jossey-Bass.

Kulm, G., & Malcom, S. (Eds.). (1994). *Science assessment in the service of reform*. Hillsdale, NJ: Erlbaum. (Original work published 1991)

Kula, S., & Bukova-Guzel, E. (2012, June). *Matematik ogretmen adaylarinin soru sormayi etkili kullanma durumlari: Limit ornegi* [Effective questioning cases of mathematics teachers: The limit case]. Paper presented in X. Ulusal Fen Bilimleri ve Matematik Egitimi Kongresi, Nigde, Turkey.

Madaus, G. F. (1988). The influence of testing on the curriculum. In L. N. Tanner (Ed.), *Critical issues in the curriculum* (pp. 83-121). Chicago, IL: University of Chicago Press.

Ministry of National Education. (1983). *Ilkokul matematik ogretmenligi programi* [Primary school mathematics teaching program]. Ankara, Turkey: Talim ve Terbiye Kurulu Baskanligi.

Ministry of National Education. (1998). *Ilkogretim matematik ogretmenligi programi* [Primary and secondary school mathematics teaching program]. Ankara, Turkey: Talim ve Terbiye Kurulu Baskanligi.

Ministry of National Education. (2005). *Ilkogretim fen ve teknoloji dersi 4 ve 5. siniflar ogretim programi* [Primary school Grades 4 and 5 science and technology teaching program]. Ankara, Turkey: Talim ve Terbiye Kurulu Başkanligi.

Ministry of National Education. (2006). *Ilkogretim fen ve teknoloji dersi 6-8. siniflar ogretim programi* [Secondary school science and technology teaching program]. Ankara, Turkey: Talim ve Terbiye Kurulu Baskanligi.

Ministry of National Education. (2008). *64 soruda ortaogretime gecis sistemi kilavuzu* [Handbook for passage to high school system in 64 questions]. Retrieved from http://bayburt.MoNE.gov.tr/teftiskurulu/duyuru2.html

Ministry of National Education. (2009). *Ilkogretim matematik dersi 6-8. siniflar ogretim programi* [Secondary school mathematics teaching program]. Ankara, Turkey: Talim ve Terbiye Kurulu Baskanligi.

National Assessment Governing Board. (2008). *Science framework for the 2009 national assessment of educational progress*. Washington, DC: Author.

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

National Science Teachers Association. (2013). *NSTA position statement: Assessment*. Retrieved from http://www.nsta.org/about/positions/assessment.aspx

Ozer, E., & Sezer, R. (2012, June). *Turkiye, Singapur ve Amerikan ders ve calisma kitaplarindaki sorularinin karsilastirmali analizi* [Comparative analysis of questions in Turkish, Singaporean, and American textbooks and workbooks]. Paper presented in X. Ulusal Fen Bilimleri ve Matematik Egitimi Kongresi, Nigde, Turkey.

Oztelli, D., Corlu, M. S., Corlu, M. A., & Capraro, R. M. (2011, April). *A comparison study of*

mathematics and science tutors and school teachers in Turkey by using international studies. Paper presented at the annual meeting of American Education Researchers Association, New Orleans, IL.

Radikal (2012). *2013'te SBS kalkiyor* [SBS will be abolished in 2013]. Retrieved from http://www.radikal.com.tr/Radikal.aspx?aType=RadikalDetayV3&ArticleID=1107681&CategoryID=86

Rico, L. (1993). Mathematics assessment in the Spanish educational system. In M. Niss (Ed.), *Cases of assessment in mathematics education: An ICMI study. New ICMI Study Series* (Vol. 1, pp. 9-20). Dordrecht, The Netherlands: Kluwer.

Rowntree, D. (1987). *Assessing students: How shall we know them?* (Rev. ed.). London, England: Kogan Page.

Senel-Coruhlu, T., Er Nas, S., & Cepni, S. (2009). *Fen ve teknoloji ogretmenlerinin alternatif olçme-degerlendirme tekniklerini kullanmada karsilastiklari problemler: Trabzon ornegi* [Problems faced by science and technology teachers in using assessment techniques: Trabzon case]. *Yuzuncu Yil Universitesi, Egitim Fakultesi Dergisi*, *4*(1), 122-141.

Shulman, L. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review, 57*(1), 1-22.

Smith, G. H., Wood, L. N., Coupland, M., Stephenson, B., Crawford, K., & Ball, G. (1996). Constructing mathematical examinations to assess a range of knowledge and skills. *International Journal of Mathematical Education in Science and Technology, 27*(1), 65-77.

Turnuklu, E. (2003). Turkiye ve Ingiltere'deki matematik ögretmenlerinin degerlendirme bicimleri [Mathematics teachers' assessment styles in Turkey and England]. *Hacettepe Universitesi Egitim Fakultesi Dergisi*, *12*(24), 108-118.

Ugurel, I., Morali, H. S., & Kesgin, S. (2012). A comparative analysis on the mathematics questions in OKS, SBS, and TIMSS under the lens of MATH taxonomy. *Gaziantep Universitesi Sosyal Bilimler Dergisi*, *11*(2), 423-444.

Wood, R. (1986). The agenda for educational measurement. In D. Nuttall (Ed.), *Assessing educational achievement.* London, England: Falmer Press.