PEABODY JOURNAL OF EDUCATION, *80*(1), 29–55


# Curriculum Policy Seen Through High-Stakes Examinations: Mathematics and Biology in a Selection of School-Leaving Examinations From the Middle East and North Africa

*Gilbert A. Valverde*

*Comparative and International Educational Policy Program*
*Department of Educational Administration and Policy Studies*
*University at Albany, State University of New York*

A study of curriculum goals set forth in school-leaving examinations in mathematics and biology from Egypt, Iran, Jordan, Lebanon, Morocco, and Tunisia benchmarked against the French baccalaureat examinations. This investigation uncovers and contrasts the goals of secondary education as they are put forward in the tests that are used in the certification of completion of secondary studies in these countries. The work takes advantage of a metric developed for the Third International Mathematics and Science Study efforts to measure educational goals for school mathematics and science. Tests from the sample of Middle Eastern and North African (MENA) countries are found to emphasize traditional content and the ability to recall and apply routine procedures and algorithms. These goals are contrasted with French school-leaving examinations and the implications of these findings for understanding and assessing curriculum policy in MENA, in a context of educational reform and the challenges of economic competiveness are considered.

Requests for reprints should be sent to Gilbert Valverde, Department of Educational Administration and Policy Studies, Education 313A, University at Albany/SUNY, Albany, NY 12222. E-mail: valverde@uamail.albany.edu

International comparisons are often invoked in educational policymaking discussions regarding curricular goals. Policymakers are frequently interested in how curriculum goals defined in their own country compare to those in other countries. There is special interest in benchmarking educational goals with those of countries they regard as their peers, their important economic competitors, key cultural referents, or those considered as prominent leaders in educational quality. Many countries are according much importance to prescribing rigorous curricula in academic disciplines as a key feature of an investment strategy in human capital intended to enhance the life chances of students and their families and the competitiveness of the nation's economy. As this attention to the intended curriculum increases among educational leaders and policymakers, it has brought about a greater interest in the international benchmarking of educational goals.

Goals for education are commonly set forth in a variety of policy instruments. Most countries present curricular intentions in curriculum guides, programs of study, textbooks, and similar documents. Comparison of these instruments has proven useful in describing differences in the opportunities afforded children to learn school subjects (Schmidt, McKnight, Valverde, Houang, & Wiley, 1997; Schmidt, Raizen, Britton, Bianchi, & Wolfe, 1997; Valverde, Bianchi, Houang, Schmidt, & Wolfe, 2002), in explaining differences in mean educational achievement across educational systems (McKnight & Valverde, 1999; Schmidt et al., 2001), in understanding the governance of curriculum (Benavot, Cha, Kamens, Meyer, & Wong, 1991; Resnick, Nolan, & Resnick, 1995; Valverde, 2000; Valverde & Schmidt, 2000), and in understanding the contribution of curriculum policies to economic growth (Benavot, 1992).

High-stakes examinations are also instruments that help scholars understand the goals of educational systems. As expressions of curricular goals, examinations are of special significance because they are formulated for presentation to students and as key referents for teachers. Students, teachers, and parents are keenly aware of the importance of demonstrating mastery of the goals as they are presented in these tests, as they recognize the consequences on the life chances of students. University entrance examinations have been studied in the past because of the evident part they play in defining criteria for the selection of students as candidates for key leadership roles in various societies (Black, 1992; Heyneman & Fägerlind, 1988; Hidano, 1988). Secondary school-leaving examinations have also been studied to understand the specific curriculum goals of educational systems and to provide practical guidance to decision makers regarding trends in how evaluation systems of the world attempt to operationalize learning goals for key actors in the educational system (Britton & Raizen, 1996; Eckstein & Noah, 1993; Kangasniemi & Takala, 1995; Madaus & Kellaghan, 1991).

Countries in the Middle East and North Africa (MENA) that have been striving to develop policies in education that enhance comprehensive efforts to reformulate investments in human capital have become interested in the benchmarking of educational goals and outcomes. Large-scale, cross-national studies of education have rarely counted countries from this region among their participants; yet, recently, an increasing number of these have participated in studies conducted by international organizations such as the International Association for the Evaluation of Educational Achievement. One such study, the 1999 replication of the Third International Mathematics and Science Study (known as TIMSS–R), had four countries from this region among its participants: Iran, Jordan, Morocco, and Tunisia. Results in that study showed mean achievement levels of eighth graders from these countries, both in mathematics and in science, to be significantly lower than the international average (Martin et al., 2000; Mullis et al., 2000).

However, disheartening as these results might have been, they were only the most recent addition to a steadily accumulating body of evidence regarding the challenges faced by educational systems in MENA. Since the 1990s, researchers and commentators have pointed out the indications of low educational quality, substantial and persistent problems of gender and urban–rural disparities in access and achievement, and the dearth of good data to guide educational reform that characterize most of the region (Anderson & Martinez, 1998; BouJaoude, 2002; Lavy & Spratt, 1997; Waterbury, 1998; World Bank, 1995).

These indications of poor performance lent momentum to efforts to substantially redirect educational priorities in the region (Belghazi, 1996; Billeh, 1996; Gargouri, 1996; Nagi, 1996). These reform efforts often set explicit goals regarding the international competiveness of educational systems in the region. Such goals have led to the increased interest in the cross-national benchmarking of student outcomes, which was reflected in the presence of four MENA countries in TIMSS–R. It also led to increased interest in other forms of benchmarking that can help provide evidence regarding structural factors that might explain the comparative achievement of students in such a way that key policy leverage points may be identified. These efforts have been substantially supported by development assistance agencies (Waterbury, 1998).

## This Study

The World Bank has been an important partner in educational policy reform efforts in MENA, and this study was commissioned among a number of other efforts as background for an Education Sector Strategy Note for the MENA region (Human Development Sector. Middle East and North Africa Region, 1998). The purpose was to take test forms from school-leaving examinations in mathematics and biology selected by the staff of the World Bank from Egypt, Iran, Jordan, Lebanon, Morocco, and Tunisia and to study them, taking advantage of techniques used in other works conducted by the first TIMSS in the study of teaching, textbooks, national curricula, and achievement conducted over the course of 3 years in 48 countries.[1] To provide an opportunity for benchmarking with a case from outside the region, data from the French baccalaureat examinations administered in Paris and Aix in 1991 and 1992 were also used.[2] The school-leaving examinations of MENA and France are both high-stakes examinations that serve a similar function in certifying the secondary school achievements of students and many of the testing systems in the region have traditionally regarded the baccalaureat as an important reference.

Data derived from these examinations were used to benchmark the intended curricula of MENA countries. This benchmarking seeks to uncover and contrast the goals of secondary education as they are put forward in the tests that are used in the certification of completion of secondary studies in these countries.

This research takes advantage of a metric developed for the TIMSS efforts to measure educational goals for school mathematics and science. The metric, known as the TIMSS Curriculum Frameworks, (Robitaille et al., 1993; Survey of Mathematics and Science Opportunities, 1992a, 1992b) measures content, performance expectations, and nondisciplinary perspectives as they are present in elements of mathematics and science curricula. The development of this international metric was grounded on extensive international consultation and validation (Schmidt et al., 1996; Schmidt & McKnight, 1995). The use of these frameworks in a detailed process of curriculum analysis has yielded comparable quantitative measures for curriculum benchmarking used in a number of investigations of curriculum policy (Schmidt, Raizen, et al., 1997; Valverde, 2000; Valverde & Schmidt, 2000).

[1]The original TIMSS, mounted in 1995, allowed for more than the measurement of achievement in mathematics and science. TIMSS also collected curriculum, textbook, and teacher data from 48 countries. The array of data collection procedures in TIMSS consisted of five separate methods, one of which is used here to elucidate educational goals as put forward in school-leaving examinations. It is a variant of *textbook analysis,* which consists of a detailed, line-by-line, page-by-page content analysis of a curriculum document.

[2]Data for these examinations were provided by the National Center for Improving Science Education, Washington, DC. The French data were coded by John Dossey, Illinois State University (mathematics), and Pinchas Tamir, Hebrew University of Jerusalem (biology). Additional information concerning the French examination data can be found in Britton and Raizen (1996).

These instruments are intended to measure three aspects of subject matter content and performances in mathematics and science, two of which were used in the work reported here.

The focus of these analyses was to study and compare goals regarding *content,* which refers to subject matter topics. These analyses also survey goals referred to as *performance expectations,* or what students are expected to do with specific contents. These comparisons are done across the test forms in the data set and are compared with portrayals of these same aspects of the French baccalaureat.

A full list of the test forms studied is presented in Table 1. The forms were selected and translated by the World Bank staff. In some cases, the translations included information regarding the year in which the tests were administered. Some of the translations also included information regarding the national subregions in which they were administered, the amount of time students were permitted to sit for the examinations, and information on the specific student population for which these particular forms were intended. When this information is available, it is indicated in Table 1.

Each of these test forms was coded using content analytic methods developed for the curriculum analysis component of the TIMSS. The unit of analysis for coding was each *test task.* This term is used instead of *test item* because many individually numbered test items on each test were made up of a series of discrete tasks. That is, items often included more than one question to which examinees were intended to respond—each one of these is termed a test task here.

Every task on each form for which a separate student response was intended was coded using categories from the TIMSS curriculum frameworks. The codes assigned to each task or intended response (often more than one) constitute the data from which analysis files were created. Two categories of codes were used: content and performance expectations. The frameworks are designed so that coding can make use of multiple categories. Any test task can be classified with as many codes as needed to represent its complexity. Each task can have a unique and potentially quite complex *signature*, or set of content and performance expectation codes assigned to it. This system is flexible, allowing simple or complex signatures as needed. Detailed descriptions regarding the use of the TIMSS curriculum frameworks in this type of content analysis of instruments of curriculum policy have previously been published (Schmidt et al., 1996; Valverde et al., 2002).

Regarding the findings reported here, it is also important to bear the following in mind: All analyses are devoted to examination of the content or performance expectations contained in each task. There were insufficient

Table 1
*Descriptions of the Examinations*

| *Examinations by Country* | *Form ID* | *Examinee Population* | *Date* | *Title* | *Administration Time* |
|---|---|---|---|---|---|
| Mathematics | | | | | |
| Egypt | EGMAL1 | Science track | 1995 | Mathematics (advanced level) | 3 hr |
| | EGMAT2 | Art tracks | 1995 | Mathematics (optional) | 2 hr |
| | EGSAT3 | Science and art tracks | 1995 | Statistics (optional) | 2 hr |
| | EGSTCI | Science track | 1995 | Calculus | 2 hr |
| Iran | IRNHM1 | 4th-grade students of physics and math division | January 1996 | Examination questions on analytic geometry | n/a |
| | IRNHM2 | 4th-grade students | May–June 1995 | Examination questions on algebra and analysis | n/a |
| | IRNHM3 | 4th-grade students | May–June 1995 | Examination questions on mathematics | n/a |
| | IRNHM4 | 4th-grade students of division of experimental science | May–June 1995 | Examination questions on modern mathematics | n/a |
| Jordan | JDMAT1 | Art track | 1995/Second term | Mathematics | 45 min |
| | JDMAT2 | Science track | 1995/Second term | Mathematics | 3 hr |
| Lebanon | LEB-1 | n/a | n/a | n/a | n/a |
| | LEB–2 | n/a | n/a | n/a | n/a |
| | LEB–3 | n/a | n/a | n/a | n/a |
| | LEB–4 | n/a | n/a | n/a | n/a |
| Morocco | MRMS01 | Mathematics section—Tetouan Academy | 1995 | Mathematics | 3 hr |
| | MRRSS1 | Science section—Rabat Academy | 1995 | Mathematics | 2 hr |
| | MRTSS1 | Science Section—Tebouan Academy | 1995 | Mathematics | 2 hr |
| Tunisia | TENEST1 | Experimental science/ technology/ technical mathematics | June 1995 | Baccalauréat examination—mathematics | |
| | TNH001 | Humanities | June 1995 | Baccalauréat examination—mathematics | |
| | TNM001 | mathematics | June 1995 | Baccalauréat examination—mathematics | |

| | | | | | |
|---|---|---|---|---|---|
| France | FRP001 | Paris | 1991 | Baccalauréat | n/a |
| | FRP001 | Paris | 1992 | Baccalauréat | n/a |
| | FRA001 | Aix | 1991 | Baccalauréat | n/a |
| | FRA02 | Aix | 1992 | Baccalauréat | n/a |
| Biology | | | | | |
| Egypt | EGYP01 | Science section | 1995 | Examination for the completion of high school general secondary exams—biology (advanced level) | 3 hr |
| | EGYPTS | Science track | 1995 | Mathematics (advanced level)[a] | |
| Iran | IRNH4S | 4th-grade students of experimental science | May–June 1995 | Examination questions of Zoa—biology | n/a |
| Jordan | J001 | Science stream | June 18, 1995 | General secondary school certificate examination—biology | 1 hr 45 min |
| Lebanon | LEB1 | n/a | n/a | n/a | n/a |
| | LEB2 | n/a | n/a | n/a | n/a |
| | LEB3 | n/a | n/a | n/a | n/a |
| | LEB4 | n/a | n/a | n/a | n/a |
| Morocco | MORC02 | Original science/ Experimental science—Tetouan Academy | June 1995 | Baccalauréat examination—natural sciences | 2 hr |
| | MORNS1 | Experimental science/ Agricultural science—Rabat Academy | n/a | Baccalauréat examination—natural sciences | 2 hr |
| Tunisia | TUNIS1 | Experimental sciences | n/a | Baccalauréat examination—natural sciences | 3 hr |
| | TUNIS2 | Mathematical–economics– management–technical studies– mathematics and technical studies– letters | June 1995 | Baccalauréat examination–natural sciences | 1 hr 30 min |
| France | FRP001 | Paris | 1991 | Baccalauréat | n/a |
| | FRP001 | Paris | 1992 | Baccalauréat | n/a |
| | FRA001 | Aix | 1991 | Baccalauréat | n/a |
| | FRA002 | Aix | 1992 | Baccalauréat | n/a |

[a] Although a mathematics test, this was also coded for science due to the large amount of physics content.

data concerning tasks that were optional or obligatory, the point value of each task or item, or the manner in which the tests were scored. As a consequence, assessments of relative emphasis in the tests are made exclusively in terms of the proportions of the total number of tasks present in the tests devoted to various contents or performance expectations. This provides an important indicator of relative emphasis, but one that is different from others that would have been possible if tasks were weighed by point value, time for solution, and so forth.

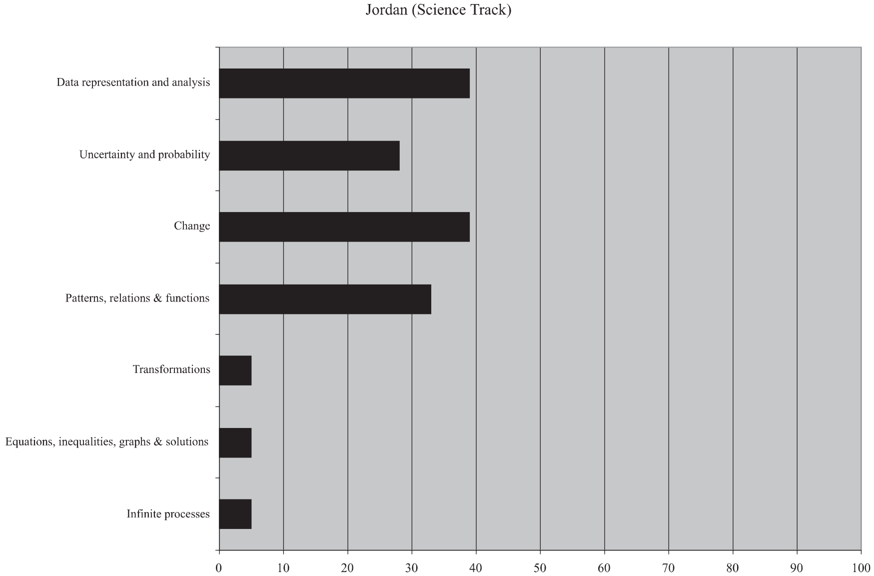
## Benchmarking Goals as Presented in Examination Forms

The benchmarking effort began with portraying the contents covered in each one of the test forms. The percentage of each test form (i.e., the percentage of the total number of test tasks in each form) devoted to each mathematics and biology topic in the TIMSS frameworks was calculated. A selection of these analyses is presented in Figure 1. It is important to note that percentages for each content might sum to more than 100 because more than one code can be assigned to any given task on the test. In addition, it must be noted that for each form, each graph is an exhaustive list of all contents contained.

Figure 1 provides an example of the variability across examination forms in terms of the mathematics topics that they assessed. Most of the Egyptian tests in the data set substantially dealt with topics related to the area of functions, relations, and equations in algebra. However, on the test intended for advanced-level students in mathematics depicted in this figure, the subject of “change”[3] in elementary analysis was dominant (this topic refers to the analysis topics of growth and decay, differentiation, integration, differential equations, and partial differentiation), in addition to the algebraic content. This contrasts with the forms in, for example, the Jordanian science track assessment, which concentrated mostly on the topic of “data representation and analysis” (tasks dealing with representing data; interpreting charts, tables, elements of statistics, etc.). A focus on algebra was also a common feature of the Moroccan forms, one of which is depicted here. This examination complemented the algebra content with a substantial percentage of tasks dealing with topics in geometry, elementary analysis, or both. Clearly, there are other notable contrasts.

Tests also presented differences in terms of the number of contents that they assessed. In the forms presented in Figure 1, we see test forms that evaluate 7, 11, and 8 specific content categories—but the range was quite

[3]A convention of this article is to enclose specific categories from the TIMSS frameworks in quotation marks (e.g., “change”).



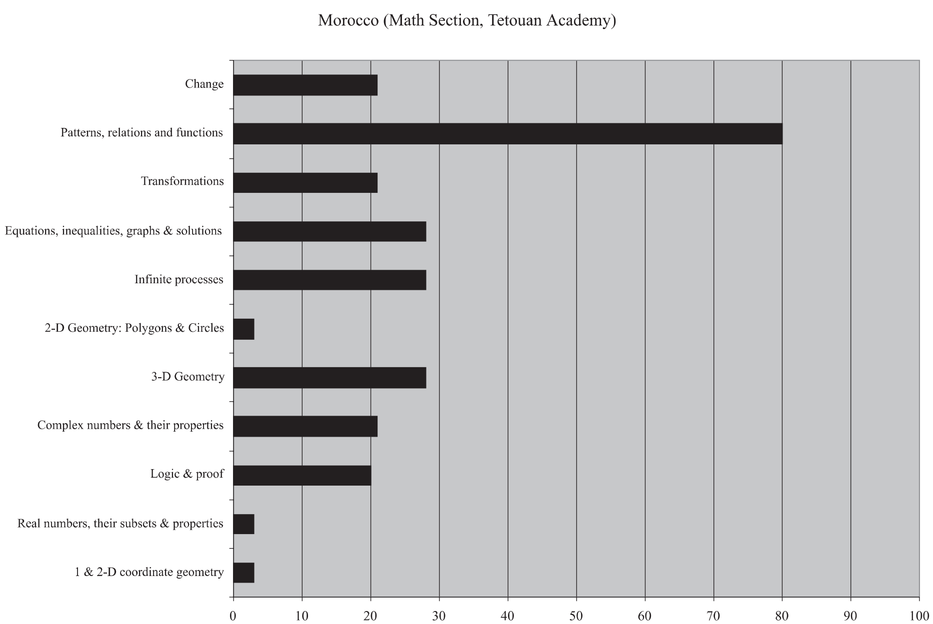




*Figure 1.* Comparison of contents of three select examination forms. Percentage of test tasks covering content is indicated by each bar. *(continued)*

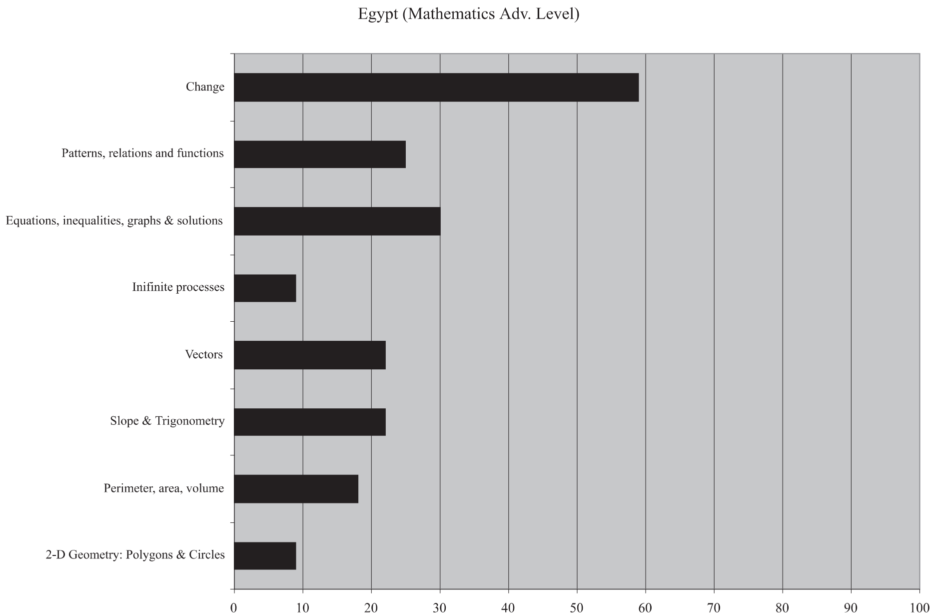



*Figure 1.* (Continued)

large across the data set, with the form assessing the fewest content areas being the arts track assessment from Jordan (3 topics) and the mathematics track baccalaureate from Tunisia (4 topics), with the tests from Lebanon standing out as particularly wide ranging (between 13 and 15 topics).

For biology, there were also differences. In separate analyses, biology tests from Tunisia and Morocco proved to have many more tasks that cover the topic of "human biology and health" than was the case for the biology tests from other countries in the sample. Two of the Lebanese examinations were very different from the others in the sample because of the preponderance of chemistry and physics topics that they contained. The presence of these topics indicated an emphasis on organic chemistry or biochemistry, which is entirely absent from the tests of other countries in the sample.

To further gauge the relative emphasis various contents received across all forms in each country, data from all forms were used to calculate a *composite curriculum core*, that is to say, lists of contents that were present in 70% or more of the tests within each country sample.[4] These topics repre-

[4]In the case of the Jordanian samples for both mathematics and biology, and the biology samples for Morocco and Tunisia, this criterion could not be fully applied because the total sample size was very small.

sent, across the forms provided for each country, another idea of which contents were considered most important for examination, as opposed to the percentage of test tasks devoted to them within each test form. In the absence of detailed information on the population of examinees and other characteristics of the curricular significance of the tests, all forms in each national sample are considered with equal weight. The topics that fulfilled the 70% criterion are listed separately for mathematics and biology, by country, in Tables 2 and 3.

In Egypt, the most commonly examined mathematics content area across the supplied forms was algebra; in Iran it was algebra and geometry. In neither of these two countries were topics in elementary analysis widely present according to this criterion, contrasting with all other country samples provided. Topics in statistics were only extensively present in the samples provided for Jordan and Tunisia. In the sample of tests provided for Lebanon, the evaluation of geometry was more common than in any of the other countries. Test forms from this country were also unique in the presence of tasks evaluating the topic of "number theory" (referring to primes, factorization, elementary number theory, etc.). The French tests contained tasks covering topics in geometry, algebra, and elementary analysis. This pattern was quite similar to the test samples in the data set from Lebanon (these tests also evaluate a number of other contents), Morocco, and Tunisia.

Table 3 shows contrasts for the biology examinations. The sample of tests from Egypt stood out in terms of the breadth of content covered. Much of the content in these tests were from the physical sciences, despite the fact that these were biology tests. The Jordanian test was distinguished by the inclusion of content related to evolution, genetics, and biological diversity (the content areas of "variation, inheritance in organisms" and "evolution, speciation, diversity"). Tunisian tests covered these same topics but also contained more topics in areas of life processes and systems enabling life functions ("energy handling," "sensing and responding," and "biochemical processes in cells"). The only topics common across Lebanese tests were from the physical sciences. The French tests clearly focused on a much more limited set of biology topics than was the case in the samples from the other countries.

Table 4 contains similar lists, calculated across all of the MENA countries in the sample. In this case, the criterion was to include in the list all topics common to 70% of the test tasks across the data set. For comparative purposes, these contents are contrasted with all of the contents present on the French forms in the sample.

In the case of mathematics, the list for MENA included four topics. These topics comprise the mathematical areas of algebra and elementary

Table 2

*Core Mathematics Topics by Country*

| *Variables* | *Egypt* | *Iran* | *Jordan* | *Lebanon* | *Morocco* | *Tunisia* | *France* |
|---|---|---|---|---|---|---|---|
| Integers, rational and real numbers | | | | | | | |
| Real numbers, their subsets, and their properties | | | | | x | | |
| Other numbers and numbers concepts | | | | | | | |
| Complex numbers and their properties | | | | | x | | x |
| Number theory | | | | x | | | |
| Systematic counting | | | x | | | | |
| Measurement | | | | | | | |
| Perimeter, area, and volume | | | | x | x | | |
| Geometry: Position, visualization, and shape | | | | | | | |
| 1-and 2-D coordinate geometry | | x | | x | x | x | x |
| 2-D geometry: Polygons and circles | | | | x | | | x |
| 3-D geometry | | | | x | | | |
| Vectors | | | | | | x | |
| Geometry: Symmetry, congruence, and similarity | | | | | | | |
| Transformations | | x | x | x | x | | x |
| Proportionality | | | | | | | |
| Slope and trigonometry | | | | x | | | |
| Functions, relations, and equations | | | | | | | |
| Pattern, relations, and functions | x | | x | x | x | x | x |
| Equations and inequalities, graphs and solutions | x | x | x | x | x | x | x |
| Data representation, probability, and statistics | | | | | | | |
| Data representation and analysis | | | x | | | | |
| Uncertainty and probability | | | x | | | | |
| Elementary analysis | | | | | | | |
| Infinite process | | | x | x | x | x | x |
| Change | | | x | x | x | x | x |
| Validation and structure | | | | | | | |
| Structuring and abstracting | | | | x | | | |

*Note.* Each country's core consists of the topics that appeared in at least 70% of tasks across its examination forms.

Table 3

*Core Science Topics in Biology Examinations by Country*

| *Variables* | *Egypt* | *Iran* | *Jordan* | *Lebanon* | *Morocco* | *Tunisia* | *France* |
|---|---|---|---|---|---|---|---|
| Earth Science | | | | | | | |
| Composition | x | | | | | | |
| History | | | | | | x | |
| Life sciences | | | | | | | |
| Plants, fungi types | x | x | x | | x | | |
| Animal types | x | x | x | | x | x | |
| Microorganism types | x | | | | | | |
| Organs, tissues | x | x | x | | x | x | |
| Cells | x | x | | | x | x | |
| Organism energy handling | x | x | x | | | | |
| Organism sensing, responding | | x | | | | x | x |
| Biochemical processes in cells | x | x | | | x | x | x |
| Life cycle of organisms | x | x | | | x | x | x |
| Reproduction of organisms | x | | | | x | x | |
| Variation, inheritance in organisms | | x | x | | x | x | x |
| Evolution, speciation, diversity | | | x | | | x | |
| Biochemistry of genetics | | x | x | | x | x | |
| Interdependence of living things | | | x | | | x | |
| Human biology and health | x | x | x | | x | x | |
| Human nutrition | x | | | | | | |

*(continued)*

Table 3 *(Continued)*

| *Variables* | *Egypt* | *Iran* | *Jordan* | *Lebanon* | *Morocco* | *Tunisia* | *France* |
|---|---|---|---|---|---|---|---|
| Human diseases | x | | | | x | x | |
| Physical sciences | | | | | | | |
| Chemical properties of matter | x | x | | | | x | |
| Atoms, molecules, ions | x | x | | | | | |
| Macromolecules, crystals | x | | | | | x | |
| Energy types, sources, conversions | x | | | x | | | |
| Heat and temperature | x | | | x | | | |
| Wave phenomena | x | | | | | | |
| Sound and vibration | x | | | | | | |
| Light | | | | x | | | |
| Electricity | x | | | x | | | |
| Magnetism | x | | | | | | |
| Chemical changes | x | x | | x | x | | |
| Organic and biochemical changes | x | x | | x | | | |
| Nuclear chemistry | x | | | | | | |
| Electrochemistry | x | | | | | | |
| Types of forces | x | | | x | | | |
| Time, space, motion | x | | | x | | | |
| Environmental and resources issues related to science | | | | | | | |
| Pollution | x | | x | | | | |
| Land, water, sea resources conservation | x | | | | | | |

*Note.* Each country's core consists of the topics that appeared in at least 70% of its tests.

Table 4

*MENA Core Content Topics Compared With France Mathematics Core Performance Expectations*

| *Topics* | *MENA Core Contents* | *France Core Contents* |
|---|---|---|
| Mathematics | | Complex numbers and their properties |
| | | Systematic counting |
| | | Perimeter, area, and volume |
| | | 1- and 2-D coordinate geometry |
| | | 2-D geometry: Basics |
| | | 2-D geometry: Polygons and circles |
| | | Vectors |
| | | Transformations |
| | Patterns, relations, and functions | Patterns, relations, and functions |
| | Equations and inequalities, graphs and solutions | Equations and inequalities, graphs and solutions |
| | Infinite processes | Infinite processes |
| | Change | Change |
| Biology | Animal types | |
| | Organs, tissues | |
| | Human biology and health | |
| | | Cells |
| | | Organism energy handling |
| | | Biochemical processes in cells |
| | | Life cycles of organism |
| | | Reproduction of organism |
| | | Variation, inheritance in organism |
| | | Evolution, speciation, diversity |
| | | Biochemistry of genetics |

*Note*. MENA = Middle Eastern and North African.

analysis as they are characterized in the TIMSS frameworks. These are also all topics that were present in the French tests from the sample. The French tests also dealt with contents related to complex numbers (including both algebraic and trigonometric forms); types of systematic counting including tree diagrams, permutations, combinations, and so forth; and a number of topics in two- and three-dimensional geometry.

In the case of biology, the composite core for the MENA sample had three topics, none of which were present in the French tests in the sample. Whereas the MENA tests gave priority to aspects of plant and animal taxonomy and morphology, the French tests focused on cells and the biochemical processes in the regulation of cell functions, translation, protein synthesis, enzymes, and similar things. The French tests also concentrated on other biological processes and topics in evolution, speciation, and diversity. This emphasis included the treatment of reproduction and the biochemistry of genetics.

Examinations do not simply intend to document the content students have mastered, but they also have the purpose of measuring whether what students are able to do with that content meets the educational system's policy goals. The intention is to elicit performances associated with content—to challenge and assess students' performances in an effort to determine whether sufficient mastery has been gained to merit certification of their secondary school experience. These expectations regarding what students are to do with content are termed *performance expectations* in the TIMSS curriculum analysis. The frameworks do not attempt to measure how test tasks or other curriculum elements might elicit identifiable cognitive processes in students. Such a strategy posited insoluble problems of both reliability and validity in coding, compounded by the cross-national character of the work for which the frameworks were designed. Instead, these framework codes identify expectations for student performances (e.g., formulating and clarifying a problem to be solved or developing a solution strategy in mathematics, or using apparatuses, gathering data, or interpreting investigational data in science are types of performance expectations that can me measured in these frameworks). The effort is to identify the performances that tasks are intended to elicit, not the specific cognitive means for producing the performances.

Figure 2 is a bar graph depicting the percentage of each nation's mathematics test forms in the sample (i.e., the percentage of the total number of tasks across tests) devoted to three major performance expectation categories. Again, it is important to note that percentages for each performance expectation also can sum to more than 100 because more than one code can be assigned to any given task on the test.

In mathematics, "performing routine procedures" was by far the most common performance expectation. This performance expectation refers to performances that demonstrate the ability to execute some procedure that the student has been exposed to and previously has practiced. It only includes performing simple procedures such as counting and standard computations, graphing, measuring, and so forth; these require very little decision making and might have been learned so well that they can be performed by rote. There were also some more demanding performance expectations present on some of the forms. One such performance expectation was in the area of formal mathematical reasoning. The aspect of mathematical reasoning most common in these tests was "justifying and proving," which required examinees to provide evidence for the validity of an action or the truth of a statement by an appeal to mathematical results or properties, or by an appeal to logic. This emphasis on justification and proof indicates an emphasis on the examination of formal mathematical structures explored for consistency and true propositions, not for the ap-

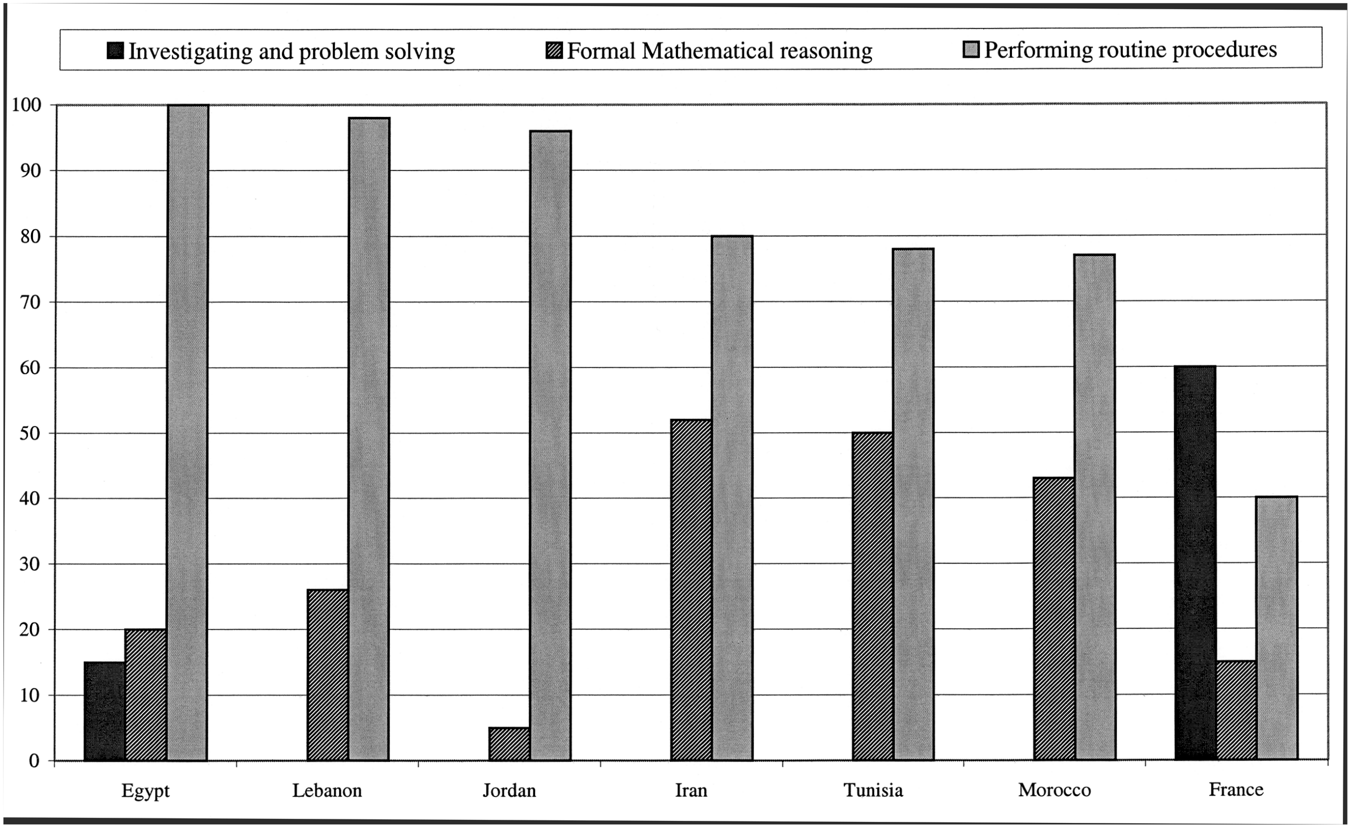



*Figure 2.* Performance expectation emphases by country.

plications of such structures to problem solving in real-world contexts. Tasks that require students to demonstrate proficiency in investigation and problem-solving skills were largely absent from these forms. Only the Egyptian tests (in fact, only one of the forms: a calculus test for the science track) had a proportion (less than 20%) of tasks requiring "formulating and clarifying problems and situations," which is one of the performance expectations in the area of investigation and problem solving, referring to formulating or clarifying a problem relating to a real-world or other concrete situation mathematically.

For biology (Figure 3), there is a clear preponderance of the "simple information" expectation under the "understanding" performance expectation category of the TIMSS science framework. This expectation refers to an examinee's ability to demonstrate understanding of vocabulary, facts, equations, and simple concepts. These test tasks thus provided opportunities for students to define, describe, and name simple concepts. Tasks that required examinees to demonstrate the ability to theorize, analyze, and solve problems are also sometimes present; particularly, this was the case for the forms provided from Lebanon. These tests incorporated a number of tasks requiring examinees to demonstrate abilities in "applying scientific principles to develop explanations" in which students were required to explain phenomena in reference to previously learned laws, theories, or principles.

To gauge the relative weight that various performance expectations received across all forms in each country, data from all forms were used to calculate *core expectations,* as was done earlier for contents. That is, lists of performance expectations that were present in 70% or more of the forms within each country sample were made.[5] These performance expectations represent, at least across the tests provided for each country, another idea of which performance expectations were most common in examinations, as opposed to the percentage of test tasks devoted to them. As was the case for content, all tests in each national sample were considered to be of equal weight. The performance expectations that fulfilled the 70% criterion are listed separately for mathematics and science, by country, in Table 5.

The only performance expectation emphasized across the test forms in the mathematics sample for Egypt was "routine procedures." All of the other country samples also stressed formal mathematical reasoning. Iran, Morocco, and Tunisia also contained a number of tasks that addressed the performance expectation category of "describing and discussing" under

[5]Again, as noted earlier, in the case of the Jordanian samples for both mathematics and biology, and the biology samples for Morocco and Tunisia, this criterion could not be applied because the total sample size was very small.

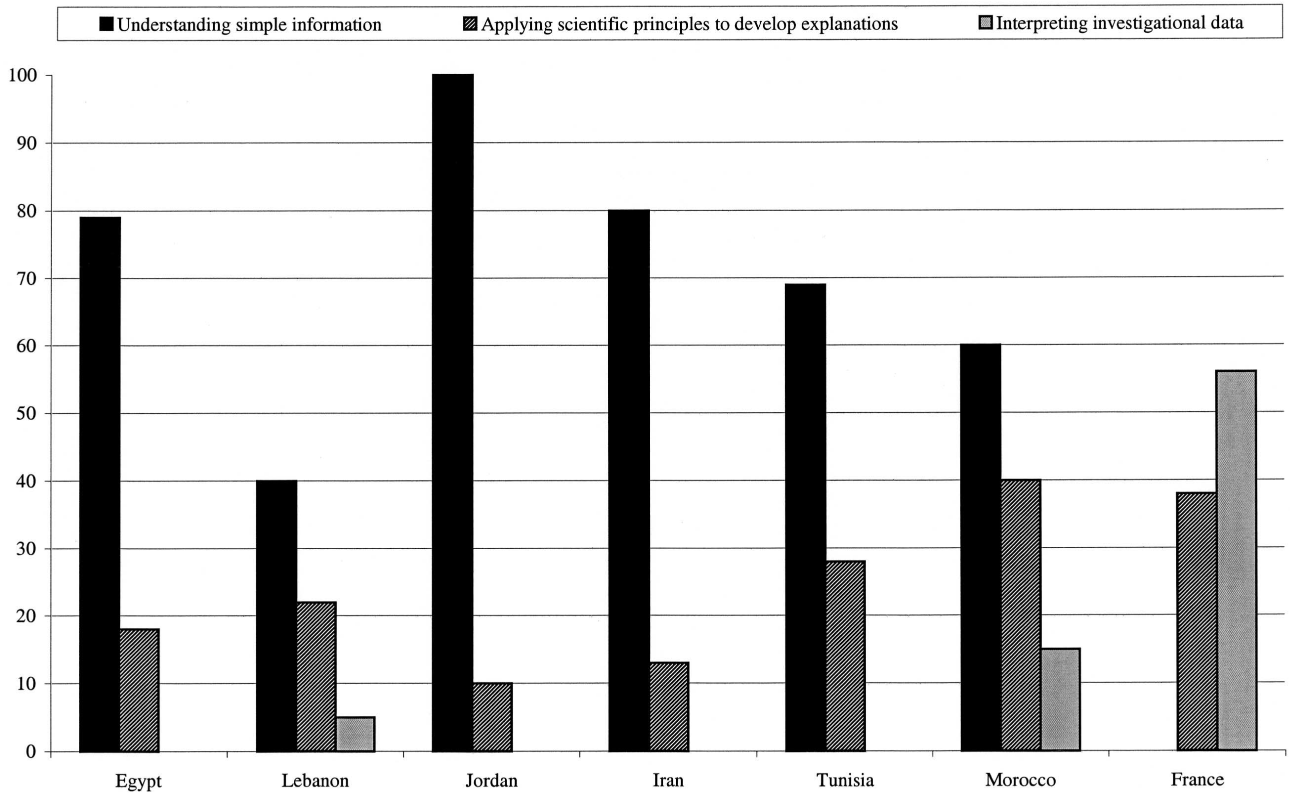



*Figure 3.* Performance expectations emphases in biology examinations.

Table 5

*Core Performance Expectations by Country*

| *Variables* | *Egypt* | *Iran* | *Jordan* | *Lebanon* | *Morocco* | *Tunisia* | *France* |
|---|---|---|---|---|---|---|---|
| Mathematics | | | | | | | |
| Knowing | | | | | | | |
| Representing | | x | x | x | x | x | |
| Recognizing equivalents | | | | | x | | |
| Recalling mathematical objects and properties | | x | x | x | | | |
| Using routine and procedure | | | | | | | |
| Performing routine procedures | x | x | x | x | x | x | x |
| Using more complex procedures | | | x | x | x | | x |
| Investigating and problem solving | | | | | | | |
| Solving | | | | | | | x |
| Verifying | | | | | | | x |
| Mathematical reasoning | | | | | | | |
| Justifying and proving | | x | x | x | x | x | |
| Communicating | | | | | | | |
| Relating representations | | | | | x | | |
| Describing and discussing | | x | | | x | x | |
| Biology | | | | | | | |
| Understanding | | | | | | | |
| Simple information | x | | x | x | x | x | |
| Complex information | x | | | x | x | x | x |
| Theorizing, analyzing, and solving problems | | | | | | | |
| Abstracting and deducing scientific principles | | | | x | x | | x |
| Applying scientific principles to solve quantitative problems | | | | x | | | |
| Applying scientific principles to develop explanations | x | | x | | x | x | x |
| Constructing, interpreting, and applying models | | | | x | | | |
| Investigating the natural world | | | | | | | |
| Making decisions | | | | | x | | |
| Interpreting investigation data | | | | | | | x |

*Note.* Each country's core consists of the performance expectations that appeared in at least 70% of its tests.

the general category of "communicating" in the mathematics framework. Tasks assessing "investigating and problem solving" were not common in any of the MENA country samples. Such tasks figure prominently in the French sample.

Table 6 contains similar lists, calculated across all of the MENA countries in the sample. In this case, the criterion was to include in the list all topics that appeared in 70% of the test forms across the MENA sample as compared with the forms in the French sample.

In the case of mathematics, the list includes four expectations from the general areas of knowing, using routine procedures, and mathematical reasoning as they are characterized in the TIMSS frameworks. Most of these were also present in the French tests from the sample. The French forms also included performance expectations in the areas of investigating and problem solving and communicating that were not common across the MENA sample considered as a whole. Certainly, these expectations for student performance in the case of the French tests challenged test takers to go beyond the recall and the application of well-practiced, routine procedures that made up the greater proportion of the MENA assessments.

Table 6

*MENA Core Performance Expectations Compared With France Biology*

| *Topics* | *MENA Sample Core Performance Expectations* | *France Core Performance Expectations* |
|---|---|---|
| Biology | | Complex information |
| | Simple information | Thematic information |
| | | Abstracting, deducing scientific principles |
| | Use science principles to explain | Use science principle to explain |
| | | Constructing and using models |
| | | Designing investigations |
| | | Interpreting investigational data |
| Mathematics | Representing | |
| | | Using equipment |
| | Performing routine procedures | Performing routine procedures |
| | Using more complex procedures | Using more complex procedures |
| | | Solving |
| | | Predicting |
| | | Verifying |
| | | Generalizing |
| | Justifying and proving | Justifying and proving |
| | | Describing and discussing |

*Note.* MENA = Middle Eastern and North African.

In the case of biology, the list for MENA includes two performance expectations. The French tests examined a larger array of performance expectations than the forms in the MENA sample considered as a whole. As was the case for mathematics examinations, the biology assessments of France challenge test takers to demonstrate a richer array of performances than was the case across MENA.

## Assessing Goals and Their Implications

The content of mathematics examinations in MENA was quite conventional. The emphasis was on algebra and calculus, with little effort to assess statistics or similar subjects. Tasks evaluating examinees' abilities in the area of problem solving were largely absent. These tests seem to indicate a conception of school mathematics as a subject largely devoted to the recognition and repetition of definitions and theorems, the performance of algorithms, and other routine procedures. There certainly were differences between the test forms and countries regarding this emphasis. Iran, Tunisia, and Morocco appear to have complemented this emphasis with aspects of formal mathematical reasoning, for example. However, the overall trend appears clear. It is particularly striking, given the emphases current in Europe and North America at the time, how these examinations made no effort to contextualize mathematics in real-world circumstances or applications.

All of the biology test samples for each of the MENA countries indicated the attempt to evaluate a very large array of topics, many more than those evaluated in tests in the French sample. The Lebanese sample was remarkable in the emphasis on physical science, despite these being biology examinations. Egyptian tests seem to have been extraordinarily comprehensive. All of the MENA tests seem to emphasize nature studies (with the exception of Lebanon) by attention to such topics as "animal types" and "organs and tissues." Clearly descriptive aspects of plant and animal taxonomy and morphology, the naming of plants and animals and their component parts, were most prominent. Less common were tasks examining major biological theories. An example is evolutionary theory, present in the one test supplied for Jordan and in the Tunisian sample. Treatment of topics related to environmental or resource conservation issues was only present to a significant degree in the examinations provided for Egypt and had a small presence in the test from Jordan. The array of performance expectations across the sampled biology tests for MENA was limited, concentrating on understanding and remembering simple facts, also with attention to the use of an understanding of scientific principals to develop explanations. The tests from Lebanon and Morocco were exceptional in the degree to which richer performance expectations were included.

The dearth of richer performance expectations is especially significant as some observers, with arguments largely supported by anecdote and little empirical evidence, have claimed that education in MENA is strongly theoretical, albeit decontextualized (Hassan, 1997). There is no evidence in these data that students were required to demonstrate much proficiency in scientific theory to fulfill requirements for their diplomas, even in the case of students taking the most advanced programs in science or mathematics. The absence of real-world contextualization in school mathematics and biology, however, was clearly reflected in these data.

Scholars that have looked at MENA have often commented on the dearth of indicators that can be used to appraise educational quality in the region. They have remarked on the sketchy and indirect evidence of how truly poor that quality is (Golladay, Berryman, Avins, & Wolff, 1998; Heyneman, 1997). This study attempts to add to the evidentiary base for assessing educational policy in MENA. This snapshot of goals in 1995 shows how secondary school goals were well aligned to the needs of static economies with workforces dominated by the large proportions employed in the state bureaucracies. As some decision makers and their international donor partners were attempting to launch reform efforts in the region, the strongest messages that students, their teachers, and their parents were receiving regarding the goal of a secondary school education were quite unambiguous. A successful secondary school student was one that could remember what he or she had been taught, one that could carefully recall the appropriate algorithms or definitions to solve types of problems in mathematics or biology with which they had been confronted in the past, and one that could understand some of the formal structures of argumentation in mathematics. As countries in MENA set out to attempt to reform their educational systems, economies, and other subsystems of their societies, the potential leaders of these reforms had graduated from secondary school systems, which did not accord a high priority to going much beyond the simple demonstration of prior knowledge.

Proficiency in applying prior knowledge to novel situations and following procedures that require extensive decision making may not be of special importance to bureaucrats and the employees of protected industries. They are, however, proficiencies of importance to a modernizing economy and to the enhancement of an individual's social and economic life chances in a changing world. Scientific and technological innovation, an important motor of modern economies, also requires educational systems to give priority to providing opportunities for students to develop proficiency in collecting, organizing, displaying, or otherwise using mathematical or science information about real-world situations.

To graduate from secondary school, students were not required to demonstrate how they could formulate and clarify problems, select or develop problem-solving strategies or data-gathering activities, or execute an ad hoc solution strategy, even one that was chosen as a result of a self-generated problem-solving strategy. At the time, such goals were embodied not only in the French school-leaving examinations used here for comparison purposes but also in those of other European and Pacific Rim countries (American Federation of Teachers and National Center for Improving Science Education, 1994a, 1994b).

What goals as embodied in school-leaving examinations can tell us regarding goals as enacted in classrooms is arguable. Certainly, there have been strong arguments regarding the stochastic effects of high-stakes tests on teaching practices that have been long advanced in the literature (Atkin, 1994; Beadie, 1999; Bloom, 1961; Division of Assessment and Accountability, 2001), and observations in MENA countries have agreed on the fact that school-leaving examinations have a determining effect on teacher's decision making (Hargreaves, 1997; Hassan, 1997). One recent empirical study conducted among mathematics and science teachers in Egypt found that teachers in that country, by their own report, regard the content of school-leaving examinations as the single most important factor in determining whether their teaching practices can be changed to address educational reform goals (Monk, Swain, Ghrist, & Riddle, 2002).

Secondary school-leaving examinations articulate a nation's aspirations regarding what youths should attain at the culmination of their pre-university studies. They specify what these young people will be held accountable for to obtain recognition of their long years of study. They also signal what will be the skills accorded the highest priority in determining places in the world of work, political leadership, or opportunities to continue on to higher education. Therefore, the importance of the examinations reviewed here is not narrowly technical–pedagogical. These high-stakes examinations are important articulations of national policy with significant social and economic consequences.

Clearly, the educational systems of these MENA countries were not preparing students to acquire the skills, dispositions, and habits of mind associated with competitive economies. All contemporary indications are that the cognitive demands of the world of work have, in fact, increased substantially since then. During their most recent meeting, the Ministers of Education of the countries of the Organization for Economic Cooperation and Development considered a report that documented the substantial "upskilling" of all work environments whether they are the occupations in information and communication technologies and services that are held to be at the vanguard of the knowledge economy, or in manufacturing, or other areas (Organization for Economic Cooperation and Development, 2001).

Already, workers in 1995 required skills in problem solving, experimentation, and communication to be competitive as globalization gathered momentum. The school-leaving examinations of MENA of that period embodied a vision of the preparation of a new generation of leaders that stressed time-honored contents and proficiencies that suggested a very different conception of the requirements of leadership in the society and the economy.

## References

American Federation of Teachers and National Center for Improving Science Education. (1994a). *What college-bound students abroad are expected to know about biology: Exams from England and Wales, France, Germany and Japan. Plus a comparative look at the United States* (Vol. 1). Washington, DC: American Federation of Teachers.

American Federation of Teachers and National Center for Improving Science Education. (1994b). *What college-bound students abroad are expected to know about chemistry and physics: Exams from England and Wales, France, Germany and Japan. Plus a comparative look at the United States* (Vol. 1). Washington, DC: American Federation of Teachers.

Anderson, R. E., & Martinez, A. (1998). Supporting private sector development in the Middle East and North Africa. In N. Shafik (Ed.), *Prospects for Middle Eastern and North African economies: From boom to bust and back?* (pp. 178–193). New York: St. Martin's Press.

Atkin, J. M. (1994). Developing world-class education standards: Some conceptual and political dilemmas. In N. Cobb (Ed.), *The future of education: Perspectives on national standards in America* (pp. 61–84). New York: College Entrance Examination Board.

Beadie, N. (1999). From student markets to credential markets: The creation of the regents examination system in New York State, 1864–1890. *History of Education Quarterly, 39*(1), 1–30.

Belghazi, S. (1996, August 6–9). *La reforme de l'enseignement au Maroc 1985–1996* [Teaching reform in Morocco 1985–1996]. Paper presented at the Conference on Education Reform in MENA, World Bank, Washington, DC

Benavot, A. (1992). Curricular content, educational expansion and economic growth. *Comparative Education Review, 36*, 150–174.

Benavot, A., Cha, Y. K., Kamens, D., Meyer, J. W., & Wong, S. Y. (1991). Knowledge for the masses: World models and national curricula, 1920–1986. *American Sociological Review, 56*, 85–100.

Billeh, V. Y. (1996). *Educational reform in Jordan: An analytical overview.* Amman, Jordan: National Center for Human Resources Development.

Black, P. (1992). *Physics examinations for university entrance.* New York: UNESCO.

Bloom, B. S. (1961). *Evaluation: A report of the seminars on examination reform organized by the University of New Delhi Grants Commission under the leadership of Dr. Benjamin Bloom.* New Delhi, India: University Grants Commission.

BouJaoude, S. (2002). Balance of scientific literacy themes in science curricula: The case of Lebanon. *International Journal of Science Education, 24*, 139–156.

Britton, E. D., & Raizen, S. A. (Eds.). (1996). *Examining the examinations: An international comparison of science and mathematics examinations for college-bound students*. Dordrecht, The Netherlands: The National Center for Improving Science Education/Kluwer Academic Publishers.

Division of Assessment and Accountability. (2001). *An examination of the relationship between higher standards and students dropping out* (Flash Research Report No. 5). New York: New York City Board of Education.

Eckstein, M., & Noah, H. (1993). *Secondary school examinations: International perspectives on policies and practice.* New Haven, CT: Yale University Press.

Gargouri, M. (1996, August 6–9). *Analysis of education reform in Tunisia.* Paper presented at the Conference on Education Reform in MENA, World Bank, Washington, DC

Golladay, F. L., Berryman, S. E., Avins, J., & Wolff, L. (1998). A human capital strategy for competing in world markets. In N. Shafik (Ed.), *Prospects for Middle Eastern and North African economies: From boom to bust and back?* (pp. 197–255). New York: St. Martin's Press.

Hargreaves, E. (1997). The diploma disease in Egypt: Learning, teaching, and the monster of the secondary leaving certificate. *Assessment in Education, 4,* 161–176.

Hassan, F. (1997). Science education in Egypt and other Arab countries in Africa and West Asia. *Frontiers, 3*(2), 1–11.

Heyneman, S. P. (1997). The quality of education in the Middle East and North Africa (MENA). *International Journal of Educational Development, 17,* 449–466.

Heyneman, S. P., & Fägerlind, I. (1988). *University examinations and standardized testing.* Washington, DC: International Bank for Reconstruction and Development.

Hidano, T. (1988). Admission to higher education in Japan. In S. P. Heyneman & I. Fägerlind (Eds.), *University examinations and standardized testing: Principles, experience and policy options* (pp. 9–25). Washington, DC: World Bank.

Human Development Sector. Middle East and North Africa Region. (1998). *Education in the Middle East and North Africa: A strategy towards learning and development.* Washington, DC: World Bank.

Kangasniemi, E., & Takala, S. (Eds.). (1995). *Pupil assessment and the role of final examinations in secondary education.* Lisse, The Netherlands: Swets & Zeitlinger.

Lavy, V., & Spratt, J. (1997). Patterns of incidence and change in Moroccan literacy. *Comparative Education Review, 41,* 120–141.

Madaus, G. F., & Kellaghan, T. (1991). *Student examination systems in the European community: Lessons for the United States.* Washington, DC: Office of Technology Assessment, United States Congress.

Martin, M. O., Mullis, V. S., Gonzalez, E. J., Gregory, K. D., Smith, T. A., Chrostowski, S. J., et al. (2000). *TIMSS 1999 International Science Report: Findings from IEA's Repeat of the Third International Science and Science Study at the Eighth Grade*. Chestnut Hill, MA: TIMSS International Study Center, Boston College.

McKnight, C. C., & Valverde, G. A. (1999). Explaining TIMSS mathematics achievement: A preliminary survey. In G. Kaiser, E. Luna, & I. Huntley (Eds.), *International comparisons in mathematics education* (pp. 48–67). London: Falmer.

Monk, M., Swain, J., Ghrist, M., & Riddle, W. (2002). Changing classroom practice in science and mathematics lessons in Egypt: Inhibitors and opportunities. *Mediterranean Journal of Educational Studies, 7*(1), 25–45.

Mullis, I. V. S., Martin, M. O., Gonzalez, E. J., Gregory, K. D., Garden, R. A., O'Connor, K. M., et al. (2000). *TIMSS 1999 international mathematics report*. *Findings from IEA's repeat of the Third International Mathematics and Science Study at the eighth grade.* Chestnut Hill, MA: Boston College, TIMSS International Study Center.

Nagi, S. Z. (1996). *Educational reform in Egypt: A report to the World Bank.* Cairo, Egypt: American University in Cairo.

Organization for Economic Cooperation and Development. (2001). *Education policy analysis: Education and skills*. Paris: Author.

Resnick, L. B., Nolan, K. J., & Resnick, D. P. (1995). Benchmarking education standards. *Educational Evaluation and Policy Analysis, 17*, 438–461.

Robitaille, D. F., Schmidt, W. H., Raizen, S. A., McKnight, C. C., Britton, E. D., & Nicol, C. (1993). *Curriculum frameworks for mathematics and science* (Vol. 1). Vancouver, Canada: Pacific Educational Press.

Schmidt, W. H., Jorde, D., Cogan, L., Barrier, E., Gonzalo, I., Moser, U., et al. (1996). *Characterizing pedagogical flow: An investigation of mathematics and science teaching in six countries.* Dordrecht, The Netherlands: Kluwer Academic.

Schmidt, W. H., & McKnight, C. C. (1995). Surveying educational opportunity in mathematics and science: An international perspective. *Educational Evaluation and Policy Analysis, 17,* 337–353.

Schmidt, W. H., McKnight, C. C., Houang, R. T., Wang, H., Wiley, D. E., Cogan, L. S., et al. (2001). *Why schools matter: A cross-national comparison of curriculum and learning.* San Francisco: Jossey-Bass.

Schmidt, W. H., McKnight, C. C., Valverde, G. A., Houang, R. T., & Wiley, D. E. (1997). *Many visions, many aims: A cross-national investigation of curricular intentions in school mathematics* (Vol. 1). Dordrecht, The Netherlands: Kluwer Academic.

Schmidt, W. H., Raizen, S. A., Britton, E. D., Bianchi, L. J., & Wolfe, R. G. (1997). *Many visions, many aims: A cross-national investigation of curricular intentions in school science* (Vol. 2). Dordrecht, The Netherlands: Kluwer Academic.

Survey of Mathematics and Science Opportunities. (1992a). *TIMSS mathematics curriculum framework* (Research Rep. No. 38). East Lansing, MI: Author.

Survey of Mathematics and Science Opportunities. (1992b). *TIMSS science curriculum framework* (Research Rep. No. 37). East Lansing, MI: Author.

Valverde, G. A. (2000). Strategic themes in curriculum policy documents: An exploration of TIMSS curriculum analysis data. *International Journal of Educational Policy Research and Practice, 1,* 133–152.

Valverde, G. A., Bianchi, L. J., Houang, R. T., Schmidt, W. H., & Wolfe, R. G. (2002). *According to the book: Using TIMSS to investigate the translation from policy to pedagogy in the world of textbooks.* Dordrecht, The Netherlands: Kluwer Academic.

Valverde, G. A., & Schmidt, W. H. (2000). Greater expectations: Learning from other nations in the quest for world-class standards in U.S. school mathematics and science. *Journal of Curriculum Studies, 32,* 651–687.

Waterbury, J. (1998). The state and economic transition in the Middle East and North Africa. In N. Shafik (Ed.), *Prospects for Middle Eastern and North African economies: From boom to bust and back?* (pp. 159–177). New York: St. Martin's Press.

World Bank. (1995). *Claiming the future: Choosing prosperity in the Middle East and North Africa.* Washington, DC: Author.