

Gulf States' Experiences with Eighth Grade Science and Mathematics

Lessons from TIMSS-2007

A paper presented at the 35th Annual Conference of the International Association for Educational Assessment, Brisbane, Australia, from 13–18 September 2009

Globalisation is nowhere more dramatic than the swiftly developing economies of the Gulf States. The Gulf Cooperation Council of six members (GCC)¹ was established in 1981 with the aim of strengthening security, economic growth, regulatory responses, and societal development in the region (GCC, n.d). Since then, the GCC countries have reached living standards and income levels equal to anywhere in the world; however, this economic revolution relied on expatriate skills. Arab societies must produce competent leaders, managers and professionals to consolidate and diversify their economies. Education, particularly in science and mathematics, is the key to this mission.

This report analyses eighth graders' science and mathematics results from TIMSS 2007² for five Gulf States: Bahrain, Qatar, Kuwait, Oman and Saudi Arabia. The United Arab Emirates was omitted, as that country was part of the benchmarking process for TIMSS. This paper discusses the adverse outcomes for these countries, which were expected, and identifies common teaching practices contributing to students' current performance levels. Recommendations to assist Gulf teaching to reach international standards are presented.

Background

A recent World Bank report (2008) reveals that despite the region's heavy investment in education, the returns were modest. The GCC countries spend an average of five per cent of their GDPs on education, spearheaded by the UAE, which allocates about 25 per cent of its federal budget on education. Nevertheless, the Gulf countries have reached almost full primary education enrolment and increased enrolment in secondary schools almost threefold between 1970 and 2003 and fivefold at the higher education level. Gender-parity for basic education is virtually complete; despite initial low levels of gender parity, indexes for secondary and higher education are not significantly different from East Asia.

In this study, international statistics on education allow GCC comparisons with a large number of indicators to pinpoint differences. The International Association for the Evaluation of Educational (IEA) in Boston USA evaluates learning outcomes for mathematics and science. TIMSS 2007, the IEA reports³, was the fourth in a cycle of international comparisons for fourth and eighth school grades to establish trends in mathematics and science achievement. To inform educational policy in the participating countries, background information is collected on the

¹ The Gulf Cooperation Council countries are presently Bahrain, Kuwait, Oman, Qatar, Saudi Arabia, and the United Arab Emirates (UAE).

² Trends in International Mathematics and Science Study 2007.

³ Accessed from the IEA site on 26 June 2009 from http://timss.bc.edu/timss2007/PDF/T07_TR_Chapter1.pdf

quantity, quality, and content of instruction. For this cycle, elements of curriculum coverage and implementation, teacher preparation, resource availability, and the use of technology were also requested.

The paper analyses the relative Gulf States' performance measures against international averages and that of the individual countries; then gender results on a similar comparison base; professional development for teachers; instructional media; class time usage; then a range of assessment tools.

Gulf States' Performance

For the TIMSS-2007 report, 67 countries participated, including 8 benchmarking participants and the GCC States. The TIMSS rankings for both mathematics and science were Advanced International Benchmark 625, High International Benchmark 550, Intermediate International Benchmark 475, and Low International Benchmark 400. The top five countries for mathematics were Chinese Taipei, Korea, Singapore, Hong Kong, and Japan; for science they were similarly Singapore, Chinese Taipei, Japan, Korea, and England. The outcomes for eighth graders from the Gulf States in mathematics and science are presented in Table 1. Table 1

GCC: Mean science and mathematics achievement

Country	Mean mathematics achievement	Mathematics placement*	Mean science achievement	Science placement*
International mean	500		500	
<u>GCC mean</u>	<u>352</u>		<u>406</u>	
Bahrain	398	35	467	26
Oman	372	41	423	36
Kuwait	354	44	418	38
Saudi Arabia	329	46	403	44
Qatar	307	48	319	47

*From 59 countries (remaining 8 were benchmarking countries)

The overall performance of the GCC students on mathematics and science tests was between the intermediate and low benchmarks. As shown in Table 1, all Gulf States performed significantly under the international mean in both subjects. In mathematics, the regional mean was 352 and the international mean, 500. Similarly in science, the regional mean was 406, compared to the international mean of 500. However, at that level, Bahrain performed well in both subjects, followed by Oman, whereas Qatar and Saudi Arabia were low in international ranking in both subjects.

Gender Differences

Gender differences in achievement of eighth graders from participating GCC countries in both subjects are presented at Table 2.

Table 2
Mean achievement in mathematics and science by gender

Mathematics	Girls		Boys		Difference in favour of girls (absolute value)
	Students %	Average point score	Students %	Average point score	
International Mean	50	453	50	448	5
<u>GCC Mean</u>	<u>51</u>	<u>369</u>	<u>49</u>	<u>335</u>	<u>34</u>
Oman	52	399	48	344	55
Qatar	50	325	50	288	37
Bahrain	49	414	51	382	32
Saudi Arabia	48	341	52	319	22
Kuwait	54	364	46	342	22

Science	Girls		Boys		Difference in favour of girls (absolute value)
	Students %	Average point score	Students %	Average point score	
International Mean	50	469	50	463	6
<u>GCC Mean</u>	<u>51</u>	<u>434</u>	<u>49</u>	<u>377</u>	<u>57</u>
Qatar	50	354	50	284	70
Bahrain	49	499	51	437	62
Oman	52	452	48	391	61
Kuwait	54	441	46	391	50
Saudi Arabia	48	426	52	383	43

Note: Standard errors in TIMSS 2007 results removed⁴.

The international mean shows significant differences for girls, five points in mathematics and six points in science. On the Gulf mean, the gender differences were higher in favour of girls, with 57 points in science and 34 in mathematics. The largest achievement differences in mathematics were in Oman, Qatar and Bahrain, and in science were similarly in Qatar, Oman and Bahrain. This outcome for eighth graders may accord with the findings of Alkhateeb (2001) who, in a study of UAE students over the 1990s, found that mathematical achievement was related to gender difference, with boys typically performing better than girls from adolescence on. Several research studies have shown that gender differences in mathematics learning are not clear during the elementary school years, but girls begin to fall behind boys during the intermediate school years, and they fall further behind during the high school years (Hedges & Nowell, 1995; Randhawa et al., 1993).

⁴ TIMSS 2007 Standard errors explained in Appendix A, accessed 29 June 2009 at http://timss.bc.edu/TIMSS2007/PDF/T07_M_IR_AppendixA.pdf. Full record available (mathematics) at http://timss.bc.edu/TIMSS2007/PDF/T07_M_IR_Chapter2.pdf and (science) http://timss.bc.edu/TIMSS2007/PDF/T07_S_IR_Chapter2.pdf

Professional Development for Teachers

The pressing requirement for teacher education programs in the GCC received considerable recent attention in media and the literature (Barber, Mourshed & Wheeler 2007 [McKinsey Quarterly]; Mukerji & Jammel 2008). Kirk and Napier (2009) note that the UAE is a significant case for issues with further teacher education, given the government's policy decision to use imported curricula and contracted expatriates to meet international objectives, whilst juxtaposing traditional Islamic society with capitalistic ventures, consumerism, and participation in the global economy. However, the UAE's case is shared throughout the GCC, as illustrated below.

For the first time, TIMSS 2007 reported on eighth grade teachers' professional development in science and mathematics, which allows comparison both internationally and regionally. These statistics are presented in Table 3.

Table 3
Teachers' professional development in mathematics and science

Professional development of teachers ¹ , by percentage of students							
Mathematics	Mathematics content %	Mathematics pedagogy/instruction %	Mathematics curriculum %	Integrating IT into mathematics %	Improving students' critical thinking/problem solving skills %	Mathematics assessment %	Overall %
International average	56	59	51	45	46	48	51
<u>GCC average</u>	<u>40</u>	<u>51</u>	<u>34</u>	<u>43</u>	<u>49</u>	<u>40</u>	<u>43</u>
Kuwait ²	45	62	30	45	69	43	49
Qatar	43	56	37	54	50	43	47
Bahrain	33	48	26	69	56	40	45
Oman	54	42	58	24	36	48	44
Saudi Arabia	26	47	19	24	34	24	29
Professional development of teachers ¹ , by percentage of students							
Science	Science content %	Science pedagogy/instruction %	Science curriculum %	Integrating IT into science %	Improving students' critical thinking/problem solving skills %	Science assessment %	Overall %
International average	58	57	51	45	46	47	51
<u>GCC average</u>	<u>49</u>	<u>55</u>	<u>36</u>	<u>44</u>	<u>45</u>	<u>45</u>	<u>46</u>
Qatar	52	68	45	55	47	48	53
Bahrain	45	50	35	68	58	50	51
Kuwait ²	53	57	33	46	47	34	45
Oman	53	45	46	20	27	61	42
Saudi Arabia	41	55	21	30	44	31	37

¹. In previous 2 years

² Data are available for between 70% and 85% of students

Analysis of professional development for teachers in mathematics shows that 43 per cent of GCC countries' students had teachers who participated in professional development during the previous two years, well below the international average (51%). Within the range of development topics, the highest differences between regional and international average student percentages were in mathematics curriculum (34% and 51% respectively) and mathematics content (40% and 56%). Within the Gulf States, Kuwait, Qatar, and Bahrain had respectively the highest percentage of students whose teachers had professional development in mathematics, with Saudi Arabia as the lowest.

In science, the percentage of students whose teachers had participated recently in professional development was generally higher than for mathematics. Nevertheless, instances of lower training for science teachers mirrored the experiences of mathematics teachers, science curriculum (36% and 51% respectively), and science content (49% and 58%) as shown in Table 5. Other types of training were within the bounds of international averages. Within the Gulf States, science pedagogy received the majority of training attention, with 55 per cent of the students having teachers that participated in this form of professional development during the past two years. The averages for the subjects 'integrating information technology into science', 'improving students' critical thinking or problem solving skills', and 'science assessment' topics were similar; 44, 45 and 45 per cent, respectively. However, within the GCC, Saudi Arabia had the lowest percentage of students whose teachers had professional development in mathematics (29%) and science (37%) respectively, whereas the other countries were similar with a range of just five per cent in mathematics.

Instructional Media

The use of textbooks in Arab countries, including the Gulf States, poses difficulties in the use of translated material, the ability to obtain sufficient quantities of prescribed and supplementary textbooks, and difficulties with changes in curricula (Al-Sadaawi 2007). To these logistical issues, Mustafa and Cullingford (2009) add an analysis of teachers' attitudes towards the generalised use of textbooks in an Arabic country in a heavily centralised education system. Primarily, teachers in this study were obliged by the Ministry of Education to follow the textbook agenda as a main source of knowledge, which reflects the position throughout the Gulf States, and therefore there was little freedom to try to change their methods of teaching. The enormous number of lessons were covered by any means and the authors found that teachers were fully dependent on 'chalk and talk' in preference to methods of enhancing student-centred learning. Further, the teachers lack the skills and access to appropriate training needed to employ different approaches to teaching. However, the authors conclude that the main problem is not caused by these factors but by the textbook itself and the attitudes it induces.

Table 4 presents teachers' reports from TIMSS 2007 on the extent of their reliance on textbooks in teaching eighth grade mathematics and science. The textbook remains the primary basis of mathematics instruction at eighth grade, with the highest student users Saudi Arabia and Qatar with 77 per cent and 70 per cent respectively. These percentages were above the international average of primary use of textbooks for mathematics (60%), whereas the other states were lower than the international average. The GCC countries were considerably higher in their usage of other means of mathematical instruction, with Kuwait having 42 per cent of its mathematics students receiving other means of learning, and that country had the lowest reliance on textbooks.

Table 4
Textbook use for teaching mathematics and science

Textbook usage, by percentage of students			
Mathematics	Use textbooks		Do not use textbooks
	Primary basis for lessons	Supplementary resource	%
	%	%	
International average	60	34	6
<u>GCC average</u>	<u>55</u>	<u>30</u>	<u>15</u>
Kuwait	23	35	42
Oman	53	41	6
Bahrain	50	29	21
Qatar	70	26	4
Saudi Arabia	77	19	4

Textbook usage, by percentage of students			
Science	Use textbooks		Do not use textbooks
	Primary basis for lessons	Supplementary resource	%
	%	%	
International average	53	40	7
<u>GCC average</u>	<u>53</u>	<u>33</u>	<u>15</u>
Oman	49	46	5
Bahrain	50	34	17
Kuwait	53	26	21
Qatar	58	25	16
Saudi Arabia*	-	-	-

* Comparable data not available.

Note: Standard errors in TIMSS 2007 results removed.

Saudi Arabian statistics were not available for science education using textbooks. The subsequent removal permitted the remaining Gulf States to group around the international average; however, there was less reliance on textbooks as a supplementary resource. Further, the region again surpassed international averages in its use of other means of science instruction, with Kuwait, and to some extent Bahrain, leading in the search for other learning strategies.

Class Time Usage

The time students spend on various activities for mathematics assists analysis of the range of learning opportunities that students undertake toward their class credits. Yushau (2006) studied the reactions of Saudi mathematics students using computers, but found no statistical difference in students' attitudes before and after the experiment to either mathematics or computers, which was not supportive of the literature findings at the time. The author postulated that the pace of change and language differences affected the outcome. However, Watson (2009) in a US study stated that, under the right circumstances, any learner can become mathematically proficient. The researcher compared the outcomes for students using both traditional and innovative mathematics curricula and found that the curriculum that helps foster students' creative thinking and problem solving, autonomy, communication skills and has the highest average will be the best in producing mathematically proficient students. This stance of varied curricula input improving outcomes was confirmed for science students by Kinniburgh and Shaw

(2009) in a US study that found that renewed attention to mathematics and reading for students allowed science results to lag. Using science texts to improve reading raised outcomes in both subjects.

Table 5 provides evidence that the time spent on activities commonly encountered in mathematics classes in the Gulf States, as reported by mathematics and science teachers, reflects traditional teaching methods.

Table 5
Mathematics and science: weekly class time division

Mathematics	Reviewing homework	Lectures/presentations	Guided problem solving	Self-driven problem solving	Reviewing previous material	Tests	Classroom administration	Other
	%	%	%	%	%	%	%	%
International average	9	16	21	22	13	10	4	5
<u>GCC average¹</u>	<u>11</u>	<u>21</u>	<u>19</u>	<u>13</u>	<u>15</u>	<u>10</u>	<u>6</u>	<u>6</u>
Bahrain ¹	11	23	18	12	15	11	6	6
Kuwait ¹	11	21	18	14	16	9	7	5
Oman ¹	11	18	20	14	15	11	6	6
Qatar ¹	11	21	20	13	14	10	6	6
Saudi Arabia ¹	12	22	17	11	15	10	7	7
Science	Reviewing homework	Lectures/presentations	Guided problem solving	Self-driven problem solving	Reviewing previous material	Tests	Classroom administration	Other
	%	%	%	%	%	%	%	%
International average	9	25	17	13	13	10	6	7
<u>GCC average^{1,2}</u>	<u>10</u>	<u>23</u>	<u>15</u>	<u>12</u>	<u>13</u>	<u>11</u>	<u>6</u>	<u>9</u>
Bahrain ²	10	24	16	10	13	12	8	7
Oman ^{1,2}	10	21	16	13	14	11	5	9
Qatar ^{1,2}	11	25	13	12	12	10	7	11

¹ Rounding error

² Data are available for between 50% and 70% of students.

Note: Science results not available for Kuwait and Saudi Arabia.

Table 5 (Mathematics) shows that, contrary to the international average of 43 per cent of class time spent on guided and self-regulated problem solving, the Gulf average was 32 per cent on these activities, with the difference shown as more lectures and work reviews. Within the Gulf States, Oman was the closest to the international average for problem solving, and it also had the least lectures. Although science class data were not available for two states, Kuwait and Saudi Arabia, the GCC trend against problem solving occurred to a lesser degree. Combined international averages for science problem-solving were 30 per cent of class time and for the Gulf region, 27 per cent. Further, the GCC countries showed less time spent on lectures than the international average, with minimal differences for the remaining activities. Qatar had a high 18 per cent in non-science classroom activities.

As an example of the Gulf States' focus on education, Qatar in 2002 commenced building a standards-based student assessment system for languages, mathematics and science (Gonzalez et al. 2009). It was designed to deliver information on school performance and feedback for teachers; and education reform progress for policymakers and to monitor the performance of the rising number of independent schools. The Qatar Comprehensive Educational Assessment (QCEA), administered in 2004 to about 88,000 students, was the first national standardised assessment platform in the Gulf region. The QCEA measures student learning and performance according to curriculum standards using a multiple-choice and open-ended question format. It is a summative assessment and is administered to K-12 students at the end of the school year. Similar systems are now in place in the other GCC countries.

This section analyses a range of assessment tools for eighth grade mathematics and science students to find GCC differences with international standards. First the preferred type of assessment is shown; next the format for tests is considered, and lastly, the type of questions asked within this format. Table 6 describes assessment instruments.

Table 6
Assessment preferences for mathematics and science

Assessment method, by preference and percentage of students									
Mathematics	Teacher's decision			Classroom tests			External achievement tests		
	Primary means %	Secondary importance %	Little/no import. %	Primary means %	Secondary importance %	Little/ no import. %	Primary means %	Secondary importance %	Little/no import. %
International average	45	42	13	65	30	5	27	38	35
<u>GCC average</u>	<u>41</u>	<u>41</u>	<u>19</u>	<u>60</u>	<u>30</u>	<u>9</u>	<u>28</u>	<u>34</u>	<u>38</u>
Bahrain	43	40	17	63	30	7	27	48	25
Kuwait	38 ¹	41	21	44 ¹	37	19	28 ¹	33	38
Oman	52	37	11	74	25	1	29	34	37
Qatar	39	41	21	61	32	7	31	29	40
Saudi Arabia	31 ¹	45	25	60	28	12	26	25	49

Assessment method, by preference and percentage of students									
Science	Teacher's decision			Classroom tests			External achievement tests		
	Primary means %	Secondary importance %	Little/no import. %	Primary means %	Secondary importance %	Little/ no import. %	Primary means %	Secondary importance %	Little/no import. %
International average	45	42	13	62	33	5	27	35	37
<u>GCC average</u>	<u>50</u>	<u>34</u>	<u>16</u>	<u>67</u>	<u>30</u>	<u>4</u>	<u>29</u>	<u>32</u>	<u>39</u>
Bahrain	58	34	8	72	27	1	37	36	26
Kuwait	62 ¹	31	7	66 ¹	29	5	27 ¹	39	34
Oman	54	31	16	71	27	2	29	34	37
Qatar	46	38	16	69	28	3	27	30	44
Saudi Arabia	29	38	33	56	37	7	24	20	56

¹Data are available for between 70% and 85% of students

Note: Differences are due to rounding

In concert with international experience (95%), Gulf mathematics teachers preferred classroom tests to monitor students' progress, with 90 per cent of students receiving tests as a high or moderate preference by the teachers for this form of assessment. All reporting states depended on classroom tests, although Kuwait had mixed assessment methods, with a similar preference for the professional judgement of the teacher (combined 79% of students). However, in a UK study, Lockwood (2007) et al. questioned the efficacy of teachers' assessments, finding large variation resulting from different mathematics achievement measures and concluding that individual teacher performance based on value-added models can be sensitive to types of assessment instruments. Interestingly, 19 per cent of Gulf mathematics students (25% for Saudi Arabia) received little or no teacher-based assessment, against 13 per cent internationally, thus the GCC practice of moving toward standards-based student assessment systems is taking effect. However, there was little support across the mathematics teachers for external achievement examinations, with 38 per cent of GCC students receiving little or no external examination, whilst the international average was 35 per cent.

The trend toward the international averages for Gulf mathematics students continued with the eighth grade science classes. There was little difference between preferences for tests, with 97 per cent of GCC students taking tests as either the primary or secondary form of assessment, against 95 per cent internationally. Teachers' professional judgement was preferred for 84 per cent of science assessment, with the international average similar (87%). This tendency was also noted by Alkharusi (2008), when studying ninth grade students and science teachers from Muscat public schools in Oman. Of the GCC countries, Bahrain was more emphatic about assessment, with higher primary percentages on the three forms of assessment than either the Gulf or international averages.

Test Items Format

A formal assessment system requires a model of student cognition and learning in the field of study, well-designed and tested assessment questions and tasks (items), and the means to make inferences regarding student performance in the given context. These elements form part of the fit for learning outcomes, instructional approach and assessment outcomes (Wilson & Carstensen 2007). However, when investigating subgroup differences on a multiple-choice and constructed-response (open-ended) test of scholastic achievement, Edwards and Arthur (2007) found the constructed-response test format a viable alternative to the traditional multiple-choice test format.

In this sub-section, this issue for the Gulf countries is pursued by means of the item format for the TIMSS 2007 averages for mathematics and science tests and examinations (Table 7).

Table 7
Item formats, mathematics and science tests/examinations

Mathematics	Constructed response		Constructed response and multiple choice		Multiple choice	
	Students %	Average Achievement	Students %	Average Achievement	Students %	Average Achievement
International average	44	453	41	448	15	440
<u>GCC average</u>	<u>15</u>	<u>344</u>	<u>73</u>	<u>353</u>	<u>12</u>	<u>342</u>
Bahrain	42	405	50	390	8	386
Oman	8	364	81	375	11	359
Kuwait	6	352	79	356	15	352
Qatar	14	302	78	312	8	286
Saudi Arabia	4	295	76	330	20	327

Science	Constructed response		Constructed response and multiple choice		Multiple choice	
	Students %	Average Achievement	Students %	Average Achievement	Students %	Average Achievement
International average	23	469	63	464	14	459
<u>GCC average</u>	<u>6</u>	<u>405</u>	<u>77</u>	<u>405</u>	<u>17</u>	<u>399</u>
Bahrain	5	440	82	471	13	459
Oman	7	428	92	423	1	n.d
Kuwait	9 ¹	419	72	414	19	423
Qatar	5	333	82	319	13	307
Saudi Arabia	2	n.d.	61	400	37	407

¹ Data are available for between 70% and 85% of students

At Table 7, GCC average mathematics test question formats for eighth year students are shown to differ markedly from the international average, relying on mixed constructed response and multiple choice items. Whilst Bahrain adopted constructed response format (42% of its eighth year mathematics students) at the level of the international average (44%) and achieved nearly 90 per cent of the international average in that format; the other countries ranged down to one-third less than the international average in their achievement results. Bahrain was higher in achievement results for both other item formats, within 13 per cent of the international average. Further, there were no significant differences among GCC results in the average achievements for the three item formats.

As Bahrain also used the mixed format for science (Table 7), there was greater uniformity among the GCC countries, and the average percentage of students using this format thus exceeded the international average. There was little regional interest in constructed response questions, where the Gulf countries (6% of students) lagged significantly behind the international average of 23 per cent. Again there was minor difference in the regional achievement averages based on the type of format used for items; however, Bahrain met the international average achievement for multiple choice questions and exceeded it (471 to 464, respectively) for the mixed format items.

Nature of Test Questions

To develop quality mathematics and science classroom assessments, aspects of capability (such as mastery of content knowledge, reasoning capability, performance skills, and disposition capabilities) require a range of assessment tools, as discussed above (Tytler et al. 2008). Archbald and Grant (2000) measured the content of middle school mathematics teachers' tests and quizzes. They found a large preponderance of single-path/single-solution problems related to number sense and number relations, thus 7th and 8th year students were involved largely in advanced arithmetic and learning mathematics as facts, procedures, and skills. Parke and Lane (2008, p.132) found that in regards to state assessment and standards for mathematics, the degree of alignment was higher for instruction than the more subjective assessment activities, which had greater variation in results.

Assessment is thus lagging knowledge learning in the literature, as it is frequently mandated to achieve consistency and thus comparability. TIMSS 2007 attempted to extricate data on this aspect by grouping assessment items. Table 8 presents information regarding cognitive aspects elicited through the mathematics and science tests for Gulf students in grade eight.

Table 8

Nature of test questions, mathematics and science

	Nature of mathematics test questions by percentage of students											
	Recall of facts & procedures			Application of mathematical procedures			Patterns & relationships			Explanations & justifications		
	Often	Not often	None	Often	Not often	None	Often	Not often	None	Often	Not often	None
	%	%	%	%	%	%	%	%	%	%	%	%
International average	52	42	6	74	24	2	22	68	10	47	47	6
<u>GCC average</u>	<u>55</u>	<u>41</u>	<u>5</u>	<u>72</u>	<u>25</u>	<u>3</u>	<u>16</u>	<u>66</u>	<u>19</u>	<u>19</u>	<u>64</u>	<u>17</u>
Bahrain	46	49	5	76	23	2	17	72	11	32	60	9
Oman	68	32	0	80	20	0	16	74	10	20	67	13
Kuwait ¹	47	39	14	52	34	14	14	59	27	15	56	30
Saudi Arabia	54	45	1	67	32	1	12	62	27	13	68	19
Qatar	58	40	3	85	15	0	19	62	19	15	69	16
	Nature of science test questions by percentage of students											
	Recall of facts & procedures			Application of knowledge & understanding			Developing hypotheses & designing investigations			Explanations & justifications		
	Often	Not often	None	Often	Not often	None	Often	Not often	None	Often	Not often	None
	%	%	%	%	%	%	%	%	%	%	%	%
International average	64	34	2	72	26	2	19	60	22	47	47	6
<u>GCC average</u>	<u>61</u>	<u>38</u>	<u>1</u>	<u>71</u>	<u>28</u>	<u>1</u>	<u>23</u>	<u>62</u>	<u>15</u>	<u>53</u>	<u>41</u>	<u>6</u>
Bahrain	72	25	3	81	17	3	29	58	13	32	60	9
Oman	50	49	1	72	28	0	9	62	29	20	67	13
Kuwait ¹	59	40	1	65	34	1	38	56	6	15	56	30
Saudi Arabia	63	37	0	73	26	0	17	70	13	13	68	19
Qatar ¹	60	39	1	63	34	3	23	66	12	15	69	16

¹ Data are available for between 70% and 85% of students
Note: Differences due to rounding.

The first part of Table 8 records teachers' responses regarding frequency of administering mathematics test questions: recall of facts and procedures, application of procedures, searching for patterns and relationships, and providing explanations/justifications. Whilst the greatest frequency of testing involved recall and applied procedural questions, the Gulf average lagged international experience in the percentages of students asked about patterns and relationships, with one-fifth of the students not receiving these test questions. This was repeated for explanations and justifications, where similar responses of 'infrequently' were recorded. Of the countries, Bahrain arguably tested students on a broader reach of mathematical principles, closer to international averages. Of note, student percentages in Kuwait and Bahrain significantly exceeded international and GCC averages for developing hypotheses. Generally, GCC assessments are based on different aspects of mathematics than the international norms.

For the science aspect of Table 8, the Gulf countries' averages were very close to the international test averages for recall of science facts and application of science knowledge. However, frequency of 'developing hypotheses' and 'explanations and justifications' differed, with student percentages from the Gulf countries exceeding the international average 23 per cent to 19 per cent respectively for frequency of test questions on developing hypotheses and 53 per cent to 47 per cent for science explanations and justifications. This suggests that the GCC countries are pursuing wider curricula for science than for mathematics, given that tests relate to class material.

Discussion

Although unique, each school subject is related to other subjects, and students' responses to mathematics and science have an effect their attitudes to school (Abu-Hilal 2000). These attitudes indirectly influence achievement although the level of aspiration has a significant direct effect on achievement. Thus this paper examines the performance of eighth graders in mathematics and science in the five GCC countries that participated in the TIMSS-2007, given that the sixth, UAE, contributed to the benchmarking process. The overall performance of the students was low, 70 per cent of the international mean for mathematics and 80 per cent for science and for the majority of the GCC countries, there was little improvement over their results from the TIMSS-2003. Factors impacting this performance are analysed, including professional development for teachers and related issues in teaching and assessing students in mathematics and science that could influence performance. Of interest, given that the more populous GCC countries have development policies in industrial and workforce sectors directed towards men more than women, girls in all Gulf States significantly outperformed boys in both subject areas. However, this may be resolved later in the high school years (Hedges & Nowell, 1995; Randhawa et al., 1993).

Curricula are frequently imported in the GCC and many teachers are expatriates; these factors impact the professional development of national staff (Kirk & Napier 2009). This study supports this view, with mathematics and science curricula and mathematics and science content especially vulnerable to inattention in the GCC. Whilst Bahrain, Kuwait and Qatar had the highest instances of students whose teachers were recently trained, Saudi Arabia had the lowest.

Due perhaps to difficulties relating to translated material, sourcing sufficient quantities of textbooks, and changes in curricula, the Gulf countries surpass the international averages of

percentages of students who use other media (Al-Sadaawi 2007). Kuwait and Bahrain have substantial proportions of students moving away from the use of textbooks for both subjects. Use of class time is also an indicator for depth of learning opportunities (Watson 2009). In this instance, Oman was closer to international mathematics averages for provision of new material and guided or self-driven problem solving; however, all Gulf States were above international mathematics averages for review and for passive presentations. Science material was close to international average for percentage of student time.

Accurate student assessment is critical in any education system, and the GCC has a priority on quality education for its member states (Gonzalez et al. 2009). Gulf mathematics teachers preferred classroom tests to monitor students' progress, although Kuwait had mixed assessment methods, with a similar preference for the professional judgement of the teacher. This form of assessment was queried by Lockwood (2007) et al., and 19 per cent of Gulf mathematics students (25% for Saudi Arabia) received little or no teacher-based assessment, against 13 per cent internationally. Science results showed fewer differences from the international norms, with Bahrain having higher primary percentages on the three forms of assessment than either the Gulf or international averages.

Quality examination items allow accuracy in inferring student performance (Wilson & Carstensen 2007). GCC average mathematics test question formats for eighth year students are shown to differ markedly from the international average, with the exception of Bahraini students, relying on mixed constructed response and multiple choice items more so than the international emphasis on constructed response. Science results were uniformly under international average, as Bahrain did not use constructed response format for primary assessment of science subjects for eighth year students. For the type of knowledge being assessed, relationships and explanations, the Gulf average was less than the international average for mathematics; however, it was greater for science.

Of the GCC countries, Qatar, Kuwait and Bahrain lead in education reform, crucial in unlocking human potential. For many students, lagging achievement as early as fourth grade is a predictor of rates of high school and college graduation, as well as lifetime earnings. It is incumbent on the GCC to gain greater investment on the countries' massive investments in education by placing greater effort on reform to allow their citizens the right to compete for local jobs that are largely the domain of a transient expatriate workforce.

References

- Abu-Hilal, M. (2000). A structural model of attitudes towards school subjects, academic aspiration and achievement. *Educational Psychology*, Vol. 20, No.1, pp. 75-84.
- Alkharusi, H. (2008). Effects of classroom assessment practices on students' achievement goals. *Educational Assessment*, Vol. 13, No. 4, pp. 243-266.
- Alkhateeb, H. M. (2001). Gender differences in mathematics achievement among high school students in the United Arab Emirates, 1991-2000, *School Science and Mathematics*, Vol. 101, No.1, pp. 5 – 9.
- Al-Sadaawi, A.S. (2007). *An investigation of performance-based assessment in science in Saudi primary schools*. Unpublished thesis, Victoria University, Melbourne, Vic. Available at <http://wallaby.vu.edu.au/adt-VVUT/public/adt-VVUT20070904.134637/>
- Archbald, D., & Grant, T. (2000) What's on the test? An analytical framework and findings from an examination of teachers' math tests. *Educational Assessment*, Vol. 6, No. 4. pp. 221-256.
- Barber, M., Mourshed, M. & Wheeler, F. (2007). Improving education in the Gulf. *McKinsey Quarterly*, Special Edition. Accessed 29 June 2009 from <http://abujoori.files.wordpress.com/2007/04/improve-gulf-education.pdf>
- Edwards, B. D., & Arthur Jr., W. (2007). An examination of factors contributing to a reduction in subgroup differences on a constructed-response paper-and-pencil test of scholastic achievement. *Journal of Applied Psychology*. Vol. 92, No.3, pp. 794-801
- Gonzalez, G., Le, V-N., Broer, M., Mariano, L.T., Froemel, J.E., Goldman, C., & DaVanzo, J. (2009). *Lessons from the field: developing and implementing the Qatar student assessment system 2002–2006*. RAND-Qatar Policy Institute, Doha.
- Gulf Cooperation Council, (n.d), *Member States of the Gulf Cooperation Council*. Accessed 26 June, 2009 from <http://www.gcc-sg.org/eng/index.php?action=GCC>
- Hedges, V., & Nowell, A. (1995). Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science*, Vol. 269, no. 5220, pp. 41-45.
- Kinniburgh, L. H. & Shaw, E. L. (2009). Science activities: using question-answer relationships to build: reading comprehension in science. *Classroom Projects and Curriculum Ideas*, Vol. 45, No. 4, pp. 19-28 .
- Kirk, D. & Napier, D. (2009). The transformation of higher education in the United Arab Emirates: issues, implications, and intercultural dimensions in J. Zajda, H. Daun and L. Saha (eds.) *Nation-building, identity and citizenship education*, Vol.3, Springer, Dordrecht, Netherlands.
- Lockwood, J., McCaffrey, D., Hamilton, L., Stecher, B., Le V-N., & Martinez F. (2007). The sensitivity of value-added teacher effect estimates to different mathematics achievement measures. *Journal of Education Measurement*, Vol.44, No.1, pp.47-67.
- Martin, M., Mullis, I., & Foy, P. (2008). *TIMSS 2007 International science report: findings from IEA's trends in international mathematics and science study at the fourth and eighth grades*. TIMSS & PIRLS International Study Center, Boston College, Chestnut Hill, MA.
- Mukerji, S. & Jammel, N. K. (2008) 'Perspectives and strategies towards collaboration in higher education in the GCC Arab States of the Gulf', *Asian Journal of Distance Education*, Vol.6, No.1, pp.76-86.
- Mullis, I., Martin, M., & Foy, P. (2008). *TIMSS 2007 International mathematics report: findings from IEA's trends in international mathematics and science study at the fourth and eighth grades*. TIMSS & PIRLS International Study Center, Boston College, Chestnut Hill, MA.

- Mustafa, M. & Cullingford, C. (2007). 'Teacher autonomy and centralised control: the case of textbooks'. *International Journal of Educational Development*, Vol. 28, No.1, pp.81-88.
- Parke, C. S. & Lane, S. (2008) Examining alignment between state performance assessment and mathematics classroom activities. *Journal of Educational Research*, Vol. 101, No. 3, pp.132-147.
- Randhawa, B. S.; Beamer, J. E.; Lundberg, I. (1993). Role of mathematics self-efficacy in the structural model of mathematics achievement. *Journal of Educational Psychology*, Vol. 85, No.1, pp. 41-48.
- Tytler, R., Groves, S., Gough, A., Darby, L., Kakkinen, C., & Doig, B. (2008, June). *Improving Middle Years Mathematics and Science*, final report for the Victorian Department of Education and Early Childhood Development, Deakin University, Melbourne, Vic.
- Watson, L. (2009, April). *An examination of two mathematics curricula and their implication on learning mathematics*, Teaching and Learning Department Capstone Project. Peabody College, Vanderbilt University, Nashville, NT.
- Wilson, M. & Carstensen, C. (2007). Assessment to improve learning in mathematics: The BEAR assessment system, in A. H. Schoenfeld, ed., *Assessing mathematical proficiency*, Volume 53, Mathematical Sciences Research Institute Publications, Berkeley, CA.
- World Bank (2008, May) *Sector brief: education in MENA*. Accessed 30 June 2009 from <http://siteresources.worldbank.org/INTMENAREGTOPEUCATION/Resources/EDUCATION-ENG-2008SM.pdf>
- Yushau, B. (2006). 'The effects of blended e-learning on mathematics and computer attitudes in pre-calculus algebra'. *Montana Mathematics Enthusiast*, Vol. 3, No.2, pp. 176-183.