Appendix B

Overview of Procedures
TIMSS 2003 Developmental Project

Process for Establishing the Mathematics Cognitive Domains for Scaling and Reporting

As explained in Chapter 1, developing reliable and valid achievement scales in the cognitive domains began with conducting a meeting of mathematics experts to examine the classification of the TIMSS 2003 items. Hosted by the IEA Secretariat in Amsterdam, 10 participants (see below) met in February 2005.

Participants in Mathematics Expert Meeting
Amsterdam, February 2005

Khattab Mohammad Abu Lebdeh – *Jordan*
Yu-Hsien Chang – *Chinese Taipei*
Tandi Clausen-May – *England*
Robert Garden – *New Zealand*
Barbara Japelj – *Slovenia*
Michael Martin – *TIMSS Study Director*
Ina Mullis – *TIMSS Study Director*
Peter Nystrom – *Sweden*
David Robitaille – *Canada*
Graham Ruddock – *England*

Based on an iterative process of discussion and classification of items, the meeting participants worked with the four cognitive domains specified in the TIMSS 2003 Framework – knowing facts and procedures; using concepts; solving routine problems; and reasoning – to devise the three cognitive domains used as the basis for this report. Essentially, the "knowing facts and procedures" and the "using concepts" domains in the TIMSS 2003 Framework were combined, and then distinctions between the combined domain and solving routine problems were clarified. Finally, distinctions were clarified between these two domains and reasoning. This process led to the three domains – knowing facts, procedures, and concepts; applying knowledge and understanding; and reasoning (see Appendix A). (For the TIMSS 2007 Framework, the participating countries suggested that these be shortened to knowing, applying, and reasoning for both mathematics and science.)

Subsequent to the Amsterdam meeting, the cognitive domains devised for the developmental project were reviewed by the TIMSS 2007 Science and Mathematics Item Review Committee (SMIRC). Hosted by the National Foundation for Educational Research in England and Wales (the institution of the IEA Chair and the TIMSS 2007 Mathematics Coordinator), this meeting was held in April 2005 in London. In particular, the SMIRC mathematics experts endorsed reporting according to the three cognitive domains and worked to further refine and clarify the description of each domain (see below for participants).

Mathematics Participants in TIMSS 2007 Science and Mathematics Item Review Committee Meeting

London, April 2005

Khattab Mohammad Abu Lebdeh – *Jordan*

Alka Arora – *TIMSS Research Associate*

Kiril Bankov – *Bulgaria*

Robert Garden – *New Zealand*

Liv Sissel Gronmo – *Norway*

Chen-yung Lin – *Chinese Taipei*

Mary Lindquist – *United States*

Ina Mullis – *TIMSS Study Director*

Graham Ruddock – *TIMSS 2007 Mathematics Coordinator*

Hanako Senuma – *Japan*

Characteristics of Items Within Cognitive Domains

IEA's TIMSS & PIRLS International Study Center (ISC) examined the spread of the items within the three domains according to item type (constructed-response or multiple-choice), content domain (algebra, geometry, etc.), and average difficulty (mean percent correct) to ensure there was sufficient coverage within each domain. As shown in Exhibit B.1, the classification resulted in a substantial number of items in each cognitive domain at both eighth grade (first page) and fourth grade (second page). Of the 194 items at the eighth grade, 65 were classified in the knowing cognitive domain, 93 in the applying cognitive domain, and 36 in the reasoning cognitive domain. Of the 159 items at the fourth grade, 58 were classified in the knowing cognitive domain, 63 in the applying cognitive domain, and 38 in the reasoning cognitive domain.

Within each cognitive domain, there was a very good spread of items in terms of item type (constructed-response or multiple-choice) at both eighth and fourth grades. Equivalent percentages of applying items were multiple-choice and constructed-response. As would

Exhibit B.1: Characteristics of Items Within Cognitive Domains

MATHEMATICS Grade 8

Number of Items by Item Type and Cognitive Domains

Item Type	Cognitive Domains			Total
	Knowing	Applying	Reasoning	
Constructed Response	11	34	21	66
Multiple Choice	54	59	15	128
Total	**65**	**93**	**36**	**194**

Percent of Score Points by Item Type and Cognitive Domains

Item Type	Cognitive Domains			Total Score Points
	Knowing	Applying	Reasoning	
Constructed Response	14%	47%	39%	85
Multiple Choice	42%	46%	12%	128
Total	**31%**	**46%**	**23%**	**213**

Number of Items by Content Domain and Cognitive Domain

Content Domain	Cognitive Domains			Total
	Knowing	Applying	Reasoning	
Number	21	31	5	57
Algebra	22	12	13	47
Measurement	7	22	2	31
Geometry	10	12	9	31
Data	5	16	7	28
Total	**65**	**93**	**36**	**194**

Percent of Score Points by Content Domain and Cognitive Domain

Content Domain	Cognitive Domains			Total Score Points
	Knowing	Applying	Reasoning	
Number	35%	55%	10%	60
Algebra	43%	23%	34%	53
Measurement	21%	73%	6%	33
Geometry	30%	36%	33%	33
Data	15%	53%	32%	34
Total	**31%**	**46%**	**23%**	**213**

Mean Percent Correct by Content Domain and Cognitive Domain

Item Difficulties (Mean Percent Correct)	Cognitive Domains			Total
	Knowing	Applying	Reasoning	
Number	50%	43%	36%	45%
Algebra	49%	45%	29%	42%
Measurement	55%	37%	41%	41%
Geometry	51%	50%	36%	46%
Data	53%	46%	34%	44%
Total	**50%**	**43%**	**33%**	**44%**

SOURCE: IEA's Trends in International Mathematics and Science Study (TIMSS) 2003

Exhibit B.1: Characteristics of Items Within Cognitive Domains

MATHEMATICS Grade 4

Number of Items by Item Type and Cognitive Domains

Item Type	Cognitive Domains			Total
	Knowing	Applying	Reasoning	
Constructed Response	17	28	24	69
Multiple Choice	41	35	14	90
Total	**58**	**63**	**38**	**159** [1]

Percent of Score Points by Item Type and Cognitive Domains

Item Type	Cognitive Domains			Total Score Points
	Knowing	Applying	Reasoning	
Constructed Response	24%	38%	38%	76
Multiple Choice	46%	39%	16%	90
Total	**36%**	**39%**	**26%**	**166**

Number of Items by Content Domain and Cognitive Domain

Content Domain	Cognitive Domains			Total
	Knowing	Applying	Reasoning	
Number	25	19	19	63
Patterns and Relationships	2	13	8	23
Measurement	10	18	4	32
Geometry	17	6	1	24
Data	4	7	6	17
Total	**58**	**63**	**38**	**159**

Percent of Score Points by Content Domain and Cognitive Domain

Content Domain	Cognitive Domains			Total Score Points
	Knowing	Applying	Reasoning	
Number	37%	29%	34%	68
Patterns and Relationships	8%	54%	38%	24
Measurement	31%	56%	13%	32
Geometry	72%	24%	4%	25
Data	24%	41%	35%	17
Total	**36%**	**39%**	**26%**	**166**

Mean Percent Correct by Content Domain and Cognitive Domain

Item Difficulties (Mean Percent Correct)	Cognitive Domains			Total
	Knowing	Applying	Reasoning	
Number	63%	56%	37%	53
Patterns and Relationships	63%	57%	36%	50
Measurement	65%	47%	39%	52
Geometry	60%	53%	43%	57
Data	69%	56%	58%	60
Total	**53%**	**53%**	**40%**	**54**

1 There were 161 items on the fourth grade mathematics assessment. Following item review, two items were deleted, and were not included in the cognitive domain scaling.

SOURCE: IEA's Trends in International Mathematics and Science Study (TIMSS) 2003

be expected, however, at both grades a relatively higher percentage of items in the knowing domain were multiple-choice, and a commensurately higher percentage of items in the reasoning domain were constructed-response. Often, the multiple-choice format is a cost-effective way to assess specific knowledge, while the constructed-response format may be required in complex problem-solving situations involving multiple strategies.

Despite some unevenness, there was good spread across content domains within each of the three cognitive domains. At eighth grade, it would have been preferable to have a higher proportion of number items in the reasoning domain (an effort is being made to address this in TIMSS 2007). That the distribution for measurement is concentrated in the applying domain makes some sense, since by eighth grade students should know about basic measurement tools and units. (In the TIMSS 2007 Framework, aspects of measurement were incorporated into the number and geometry content domains because there is little emphasis on measurement in eighth-grade mathematics curricula around the world).

Because algebra is generally not taught as a formal subject in primary school, only introductory concepts about patterns and relationships are assessed at the fourth grade. As such, a higher proportion of patterns and relationship items in the knowing category would have been preferable at the fourth grade. (In the TIMSS 2007 Framework, the patterns and relationships content domain has been incorporated into the number content domain.) Also, a higher proportion of measurement items in the reasoning domain would have been better. The low coverage of geometry in the reasoning domain is understandable, since this is a subject little emphasized at the fourth grade. (In the TIMSS 2007 Framework, the geometry content domain, now called geometric shapes and measures, has been recast to better describe the fourth-grade curricula of participating countries.)

Finally, Exhibit B.1 also shows a good range in item difficulty (mean percentage correct) internationally, on average, within each of the three cognitive domains. As would be anticipated, at both grades there was the same overall pattern, with the reasoning items the most difficult. Essentially, this pattern of the items in the reasoning domain being more difficult than the knowing or applying items was consistent across the content domains at both grades.

Constructing Achievement Scales in the Mathematics Cognitive Domains

The scaling methodology was identical to that used to report mathematics achievement results and achievement in the mathematics content domains in the TIMSS 2003 International Report. It is described in detail in Gonzalez, Galia, and Li (2004).

The TIMSS 2003 goals of broad coverage of the mathematics and science curriculum and of measuring trends across assessments necessitated a complex matrix-sampling booklet design, with individual students responding to a subset of the mathematics and science items in the assessment but not the entire assessment item pool. Given the complexities of the data collection and the need to have student scores on the entire assessment for analysis and reporting purposes, TIMSS 2003 relied on Item Response Theory (IRT) scaling to describe student achievement on the assessment and to provide accurate measures of trends from previous assessments. The TIMSS IRT scaling approach used multiple imputation, or "plausible values" methodology, to obtain proficiency scores in mathematics and science for all students, even though each student responded to only a part of the assessment item pool. To enhance the reliability of the student scores, the TIMSS scaling combined student responses to the items they were administered with information about students' backgrounds, a process known as "conditioning."

Using routine TIMSS procedures, three distinct IRT scaling models, depending on item type and scoring procedure, were used in constructing achievement scales for the mathematics cognitive domains. Each scaling model is a "latent variable" model that describes the probability that a student will respond in a specific way to an item in terms of the respondent's proficiency, which is an unobserved or "latent" trait, and various characteristics (or "parameters") of the item. A three-parameter model was used with multiple-choice items, which were scored as correct or incorrect, and a two-parameter model for constructed-response items with just two response options, which also were scored as correct or incorrect. Since each of these item types has just two response categories, they are known as dichotomous items. A partial credit model was used with polytomous constructed-response items, i.e., those with more than two score points.

Item Calibration

The first step in constructing the cognitive domain scales was to estimate the IRT model parameters for each item on each of the cognitive domain scales. This procedure, known as item calibration, was implemented using the PARSCALE software applied to a self-weighting random sample of 1000 students from each country's TIMSS 2003 student sample. Using student samples of equal size ensured that the data from each country contributed equally to the item calibration, while keeping the amount of data to be analyzed to a reasonable size.

At the fourth and eighth grades, separate calibrations were conducted for each of the three mathematics cognitive domains: knowing, applying, and reasoning (abbreviated labels). At the eighth grade, the calibrations were based on 46,000 student records; 1,000 from each of the 46 countries that participated in the 2003 assessment. At the fourth grade, the calibrations were based on 26,000 student records, 1,000 from each of the 26 countries that participated in the 2003 assessment at the fourth grade.

Evaluating the Fit of the IRT Models

After the calibrations were completed, checks were performed to verify that the item parameters obtained from PARSCALE for the three cognitive scales were a good fit for the data. An item is said to fit the IRT model when the empirical distribution of student responses (i.e., the proportion of correct student responses at various levels of student proficiency) closely matches the theoretical item response curve constructed from the estimated item parameters. For every item at both grades, the empirical and theoretical distributions were plotted and compared.

Generating IRT Proficiency Scores

Following item calibration, Educational Testing Service's MGROUP program was used to generate the IRT proficiency scores for the cognitive domain scales. This program takes as input the students' responses to the items they were given, the item parameters estimated at the calibration stage, and the conditioning variables derived from student background variables, and generates as output the plausible values that represent student proficiency.

Plausible values generated by the conditioning program are initially on the same scale as the item parameters used to estimate them. This scale metric is generally not useful for reporting purposes since it is somewhat arbitrary, ranges between approximately –3 and +3, and has a mean of zero across all countries. The plausible values for each cognitive domain scale were transformed to the same metric as the overall mathematics scale in 2003, as was done for the content domain scaling in 2003. Thus, for the eighth grade, each of the three cognitive domain scales were set to have a mean of 467 and standard deviation of 100, and for the fourth grade, a mean of 495 and standard deviation of 100.

Reliability

Exhibit B.2 displays the reliability coefficient for each country for the mathematics test overall and for the knowing, applying, and reasoning cognitive domains. The first page shows the reliabilities for the eighth grade and the second page shows the reliabilities for the fourth grade. Reliability was measured as the ratio of sampling variance to sampling variance plus imputation variance. This approach is more suitable for multiple-matrix-sampling designs where students respond to relatively few items than classical reliability methods (such as the well-known Kuder-Richardson formulas) that are affected by the number of items taken by the student. Reliability coefficients greater than .80 are generally considered acceptable for such designs.

At both grade levels, despite some variation, reliabilities generally were high for most countries. The international median (the median of the reliability coefficients for all countries) was .96 at the eighth grade and .97 at the fourth grade for the overall mathematics assessment. At the eighth grade, the median reliabilities for the cognitive domains were .93 for knowing, .96 for applying, and .88 for reasoning. At the fourth grade, they were .92 for knowing, .93 for applying, and .91 for reasoning.

Exhibit B.2: Reliabilities of Overall Mathematics and Cognitive Domains

MATHEMATICS Grade 8

Countries	Reliabilities of Overall Mathematics and Cognitive Domains			
	Overall	Knowing	Applying	Reasoning
Armenia	0.97	0.96	0.92	0.93
Australia	0.99	0.97	0.99	0.97
Bahrain	0.83	0.51	0.72	0.46
Belgium (Flemish)	0.95	0.98	0.97	0.89
Botswana	0.72	0.67	0.73	0.56
Bulgaria	0.96	0.99	0.95	0.88
Chile	0.88	0.83	0.96	0.67
Chinese Taipei	0.96	0.99	0.96	0.92
Cyprus	0.79	0.33	0.91	0.90
Egypt	0.95	0.96	0.98	0.85
England	0.99	0.93	0.99	0.88
Estonia	0.96	0.95	0.96	0.87
Ghana	0.87	0.80	0.92	0.80
Hong Kong, SAR	0.95	0.88	0.97	1.00
Hungary	0.99	0.96	0.95	0.91
Indonesia	0.98	0.98	0.98	0.93
Iran, Islamic Rep. of	0.97	0.79	0.86	0.48
Israel	0.96	0.92	0.97	0.96
Italy	0.97	0.97	0.97	0.95
Japan	0.92	0.92	0.96	0.91
Jordan	0.99	0.93	0.97	0.72
Korea, Rep. of	0.71	0.82	0.67	0.79
Latvia	0.97	0.97	0.96	0.88
Lebanon	0.97	0.95	0.96	0.86
Lithuania	0.99	0.78	0.96	0.97
Macedonia, Rep. of	0.98	0.78	0.92	0.86
Malaysia	0.98	0.98	0.95	0.91
Moldova, Rep. of	0.98	0.88	0.96	0.93
Morocco	0.77	0.81	0.56	0.50
Netherlands	0.98	0.98	1.00	0.87
New Zealand	0.97	0.97	0.98	0.96
Norway	0.94	0.95	0.82	0.84
Palestinian Nat'l Auth.	0.96	0.90	0.85	0.95
Philippines	0.98	0.95	0.98	0.87
Romania	0.99	0.96	0.99	0.92
Russian Federation	0.91	0.93	0.96	0.82
Saudi Arabia	0.95	0.84	0.98	0.94
Scotland	0.93	0.92	0.97	0.99
Serbia	0.91	0.83	0.86	0.74
Singapore	0.96	0.98	0.97	0.97
Slovak Republic	0.96	0.97	0.92	0.92
Slovenia	0.91	0.68	0.78	0.74
South Africa	0.94	0.94	0.91	0.97
Sweden	0.95	0.93	0.91	0.84
Tunisia	0.94	0.72	0.83	0.52
United States	0.98	0.99	0.99	0.87
International Median	0.96	0.93	0.96	0.88
Benchmark Participants:				
Basque Country, Spain	0.86	0.76	0.93	0.89
Ontario Province, Can.	0.95	0.89	0.97	0.97
Quebec Province, Can.	0.98	0.96	0.98	0.87
Indiana State, US	1.00	0.97	0.97	0.84

SOURCE: IEA's Trends in International Mathematics and Science Study (TIMSS) 2003

Exhibit B.2: Reliabilities of Overall Mathematics and Cognitive Domains

MATHEMATICS Grade 4

Countries	Reliabilities of Overall Mathematics and Cognitive Domains			
	Overall	Knowing	Applying	Reasoning
Armenia	0.99	0.94	0.99	0.97
Australia	0.97	0.98	0.98	0.91
Belgium (Flemish)	0.99	0.83	0.76	0.91
Chinese Taipei	0.94	0.88	0.89	0.79
Cyprus	0.98	0.89	0.91	0.83
England	0.98	0.84	0.95	0.97
Hong Kong, SAR	0.95	0.95	0.93	0.91
Hungary	0.95	0.91	0.96	0.93
Iran, Islamic Rep. of	0.93	0.83	0.95	0.95
Italy	0.99	0.98	0.93	0.96
Japan	0.90	0.71	0.74	0.77
Latvia	0.97	0.97	0.87	0.90
Lithuania	0.99	0.99	0.88	0.94
Moldova, Rep. of	0.98	0.92	0.99	0.94
Morocco	0.95	0.93	0.98	0.91
Netherlands	0.91	0.92	0.79	0.59
New Zealand	0.99	0.88	0.87	0.80
Norway	0.97	0.89	0.92	0.87
Philippines	0.99	0.97	0.98	0.93
Russian Federation	0.99	0.95	1.00	0.99
Scotland	0.95	0.92	0.91	0.87
Singapore	0.99	0.95	1.00	1.00
Slovenia	0.97	0.80	0.86	0.98
Tunisia	0.93	0.88	0.93	0.91
United States	0.99	0.96	0.85	0.93
International Median	0.97	0.92	0.93	0.91
Benchmark Participants:				
Ontario Province, Can.	0.98	0.95	0.99	0.98
Quebec Province, Can.	0.94	0.84	0.85	0.91
Indiana State, US	1.00	0.79	0.93	0.75

SOURCE: IEA's Trends in International Mathematics and Science Study (TIMSS) 2003

Correlations

Exhibit B.3 presents the Pearson correlation coefficient indicating the linear relationship between achievement in each cognitive domain and achievement on the overall mathematics assessment for each of the TIMSS 2003 countries. The first page shows the correlations for the eighth grade and the second page the correlations for the fourth grade. All of the correlations are substantial, indicating that high performance in each of the three cognitive domains is likely to be associated with high performance on the mathematics assessment overall. This means proficiency in each of the domains is an important contributor to mathematics proficiency in general.

At eighth grade, correlations were highest for knowing and applying, with a median correlation with overall mathematics achievement of .88 in each case. This means that students with high scores in these domains were equally likely to have high scores on mathematics overall. The correlation between reasoning and overall achievement was generally lower, with a median correlation of .77 (consistent with the somewhat lower reliability of the reasoning scale). This means that students with high scores in the reasoning domain also were likely to have high scores on mathematics overall, but somewhat less likely than students with high scores in the knowing or applying domains.

At the fourth grade, correlations between achievement in the cognitive domains and overall mathematics were more uniform, with correlations of .84 for the knowing domain, .86 for the applying domain, and .83 for the reasoning domain. This means that students with high scores in any one of the three cognitive domains were equally likely to have high scores on mathematics overall.

Correlations between the three cognitive scales are presented in Exhibit B.4 for the eighth grade (first page) and for the fourth grade (second page). As would be expected of cognitive domains within a single subject area, mathematics, country-level correlations at the eighth grade were generally moderate to high, with international

Exhibit B.3: Correlations of Mathematics Cognitive Domains with Overall Mathematics

MATHEMATICS Grade 8

Countries	Pearson Correlations of Mathematics Cognitive Domains with Overall Mathematics		
	Knowing	Applying	Reasoning
Armenia	0.85	0.87	0.79
Australia	0.89	0.90	0.81
Bahrain	0.81	0.81	0.69
Belgium (Flemish)	0.91	0.91	0.83
Botswana	0.75	0.76	0.64
Bulgaria	0.88	0.88	0.76
Chile	0.85	0.85	0.71
Chinese Taipei	0.92	0.93	0.86
Cyprus	0.86	0.87	0.76
Egypt	0.84	0.84	0.73
England	0.89	0.90	0.79
Estonia	0.88	0.89	0.79
Ghana	0.65	0.68	0.54
Hong Kong, SAR	0.89	0.90	0.82
Hungary	0.90	0.91	0.82
Indonesia	0.86	0.86	0.71
Iran, Islamic Rep. of	0.81	0.81	0.70
Israel	0.89	0.89	0.79
Italy	0.88	0.88	0.77
Japan	0.90	0.90	0.83
Jordan	0.86	0.86	0.76
Korea, Rep. of	0.90	0.91	0.83
Latvia	0.87	0.88	0.79
Lebanon	0.81	0.82	0.63
Lithuania	0.88	0.89	0.78
Macedonia, Rep. of	0.86	0.87	0.75
Malaysia	0.89	0.90	0.78
Moldova, Rep. of	0.85	0.85	0.74
Morocco	0.70	0.70	0.56
Netherlands	0.90	0.91	0.81
New Zealand	0.88	0.89	0.77
Norway	0.85	0.86	0.77
Palestinian Nat'l Auth.	0.82	0.83	0.72
Philippines	0.82	0.83	0.71
Romania	0.89	0.90	0.78
Russian Federation	0.88	0.89	0.76
Saudi Arabia	0.70	0.71	0.58
Scotland	0.89	0.90	0.79
Serbia	0.89	0.89	0.80
Singapore	0.92	0.92	0.85
Slovak Republic	0.89	0.90	0.82
Slovenia	0.87	0.88	0.75
South Africa	0.82	0.83	0.70
Sweden	0.87	0.87	0.75
Tunisia	0.74	0.75	0.57
United States	0.91	0.91	0.82
International Median	0.88	0.88	0.77
Benchmark Participants:			
Basque Country, Spain	0.83	0.84	0.71
Ontario Province, Can.	0.86	0.88	0.74
Quebec Province, Can.	0.84	0.87	0.74
Indiana State, US	0.81	0.82	0.73

SOURCE: IEA's Trends in International Mathematics and Science Study (TIMSS) 2003

Exhibit B.3: Correlations of Mathematics Cognitive Domains with Overall Mathematics

MATHEMATICS Grade 4

Countries	Pearson Correlations of Mathematics Cognitive Domains with Overall Mathematics		
	Knowing	Applying	Reasoning
Armenia	0.81	0.84	0.77
Australia	0.86	0.87	0.84
Belgium (Flemish)	0.80	0.83	0.78
Chinese Taipei	0.82	0.84	0.81
Cyprus	0.85	0.88	0.84
England	0.87	0.89	0.85
Hong Kong, SAR	0.81	0.84	0.81
Hungary	0.85	0.88	0.83
Iran, Islamic Rep. of	0.78	0.80	0.71
Italy	0.86	0.88	0.83
Japan	0.83	0.86	0.82
Latvia	0.84	0.87	0.83
Lithuania	0.85	0.87	0.83
Moldova, Rep. of	0.85	0.88	0.83
Morocco	0.72	0.74	0.63
Netherlands	0.77	0.82	0.76
New Zealand	0.87	0.88	0.86
Norway	0.82	0.85	0.79
Philippines	0.82	0.83	0.77
Russian Federation	0.85	0.88	0.85
Scotland	0.84	0.86	0.81
Singapore	0.85	0.89	0.87
Slovenia	0.84	0.86	0.83
Tunisia	0.75	0.77	0.66
United States	0.85	0.88	0.85
International Median	0.84	0.86	0.83
Benchmark Participants:			
Ontario Province, Can.	0.84	0.86	0.83
Quebec Province, Can.	0.82	0.84	0.80
Indiana State, US	0.77	0.79	0.75

SOURCE: IEA's Trends in International Mathematics and Science Study (TIMSS) 2003

Exhibit B.4: Correlations of Mathematics Cognitive Domains

MATHEMATICS Grade 8

Countries	Pearson Correlations for Mathematics Cognitive Domains		
	Knowing Applying	Knowing Reasoning	Applying Reasoning
Armenia	0.91	0.79	0.86
Australia	0.95	0.83	0.84
Bahrain	0.91	0.75	0.78
Belgium (Flemish)	0.95	0.85	0.85
Botswana	0.89	0.74	0.76
Bulgaria	0.95	0.80	0.80
Chile	0.94	0.78	0.76
Chinese Taipei	0.97	0.89	0.88
Cyprus	0.95	0.82	0.80
Egypt	0.95	0.81	0.82
England	0.95	0.80	0.82
Estonia	0.94	0.82	0.82
Ghana	0.74	0.60	0.62
Hong Kong, SAR	0.95	0.85	0.84
Hungary	0.95	0.85	0.84
Indonesia	0.94	0.76	0.77
Iran, Islamic Rep. of	0.92	0.77	0.78
Israel	0.95	0.83	0.81
Italy	0.95	0.80	0.79
Japan	0.96	0.86	0.86
Jordan	0.94	0.84	0.82
Korea, Rep. of	0.96	0.86	0.85
Latvia	0.94	0.82	0.82
Lebanon	0.91	0.67	0.68
Lithuania	0.95	0.81	0.81
Macedonia, Rep. of	0.95	0.78	0.82
Malaysia	0.96	0.81	0.83
Moldova, Rep. of	0.95	0.79	0.79
Morocco	0.86	0.65	0.71
Netherlands	0.94	0.81	0.84
New Zealand	0.94	0.79	0.79
Norway	0.93	0.82	0.81
Palestinian Nat'l Auth.	0.93	0.81	0.79
Philippines	0.93	0.79	0.80
Romania	0.95	0.82	0.81
Russian Federation	0.94	0.77	0.78
Saudi Arabia	0.80	0.66	0.65
Scotland	0.95	0.82	0.81
Serbia	0.95	0.84	0.84
Singapore	0.96	0.88	0.86
Slovak Republic	0.95	0.85	0.85
Slovenia	0.94	0.77	0.78
South Africa	0.89	0.76	0.78
Sweden	0.93	0.76	0.74
Tunisia	0.87	0.65	0.68
United States	0.97	0.86	0.85
International Median	0.95	0.81	0.81
Benchmark Participants:			
Basque Country, Spain	0.93	0.78	0.75
Ontario Province, Can.	0.92	0.75	0.76
Quebec Province, Can.	0.92	0.75	0.76
Indiana State, US	0.93	0.82	0.80

SOURCE: IEA's Trends in International Mathematics and Science Study (TIMSS) 2003

Exhibit B.4: Correlations of Mathematics Cognitive Domains

MATHEMATICS Grade 4

Countries	Pearson Correlations for Mathematics Cognitive Domains		
	Knowing Applying	Knowing Reasoning	Applying Reasoning
Armenia	0.84	0.74	0.86
Australia	0.92	0.89	0.91
Belgium (Flemish)	0.89	0.80	0.84
Chinese Taipei	0.92	0.87	0.91
Cyprus	0.92	0.87	0.91
England	0.94	0.89	0.91
Hong Kong, SAR	0.91	0.85	0.90
Hungary	0.90	0.82	0.89
Iran, Islamic Rep. of	0.86	0.73	0.81
Italy	0.92	0.84	0.88
Japan	0.91	0.84	0.89
Latvia	0.91	0.85	0.88
Lithuania	0.93	0.86	0.90
Moldova, Rep. of	0.89	0.82	0.89
Morocco	0.80	0.63	0.74
Netherlands	0.87	0.80	0.85
New Zealand	0.93	0.88	0.90
Norway	0.92	0.80	0.86
Philippines	0.90	0.83	0.86
Russian Federation	0.88	0.85	0.90
Scotland	0.91	0.85	0.87
Singapore	0.92	0.86	0.94
Slovenia	0.91	0.87	0.92
Tunisia	0.80	0.69	0.73
United States	0.93	0.88	0.92
International Median	0.91	0.85	0.89
Benchmark Participants:			
Ontario Province, Can.	0.91	0.87	0.90
Quebec Province, Can.	0.91	0.83	0.86
Indiana State, US	0.90	0.83	0.88

SOURCE: IEA's Trends in International Mathematics and Science Study (TIMSS) 2003

medians of .95 between the knowing and applying domains, .81 between the knowing and reasoning domains, and .81 between the applying and reasoning domains. The highest correlation was between the knowing and applying domains, which makes sense considering that these were the two domains with the highest correlation with mathematics achievement overall.

At the fourth grade, country-level correlations between the cognitive domains also were high, with international medians of .91 between the knowing and applying domains, .85 between the knowing and reasoning domains, and .89 between the applying and reasoning domains. The relatively large correlations between the cognitive domain scales show that student performance in the cognitive domains is not independent, and that high-scoring students on one scale are likely also to be high scorers on another. Despite the high correlations, however, there is scope for interesting average score differences between countries on the three cognitive scales.