Findings from a Developmental Project

Lynch School of Education, Boston College
TIMSS & PIRLS International Study

International Association
for the Evaluation of
Educational Achievement

TIMSS 2003 International Mathematics Report / by Ina V.S. Mullis, Michael O. Martin, Pierre Foy.

Publisher:
TIMSS & PIRLS International Study Center,
Lynch School of Education, Boston College

Library of Congress Catalog Card Number: 2005900696

ISBN: 1-889938-38-6

For more information about TIMSS contact:

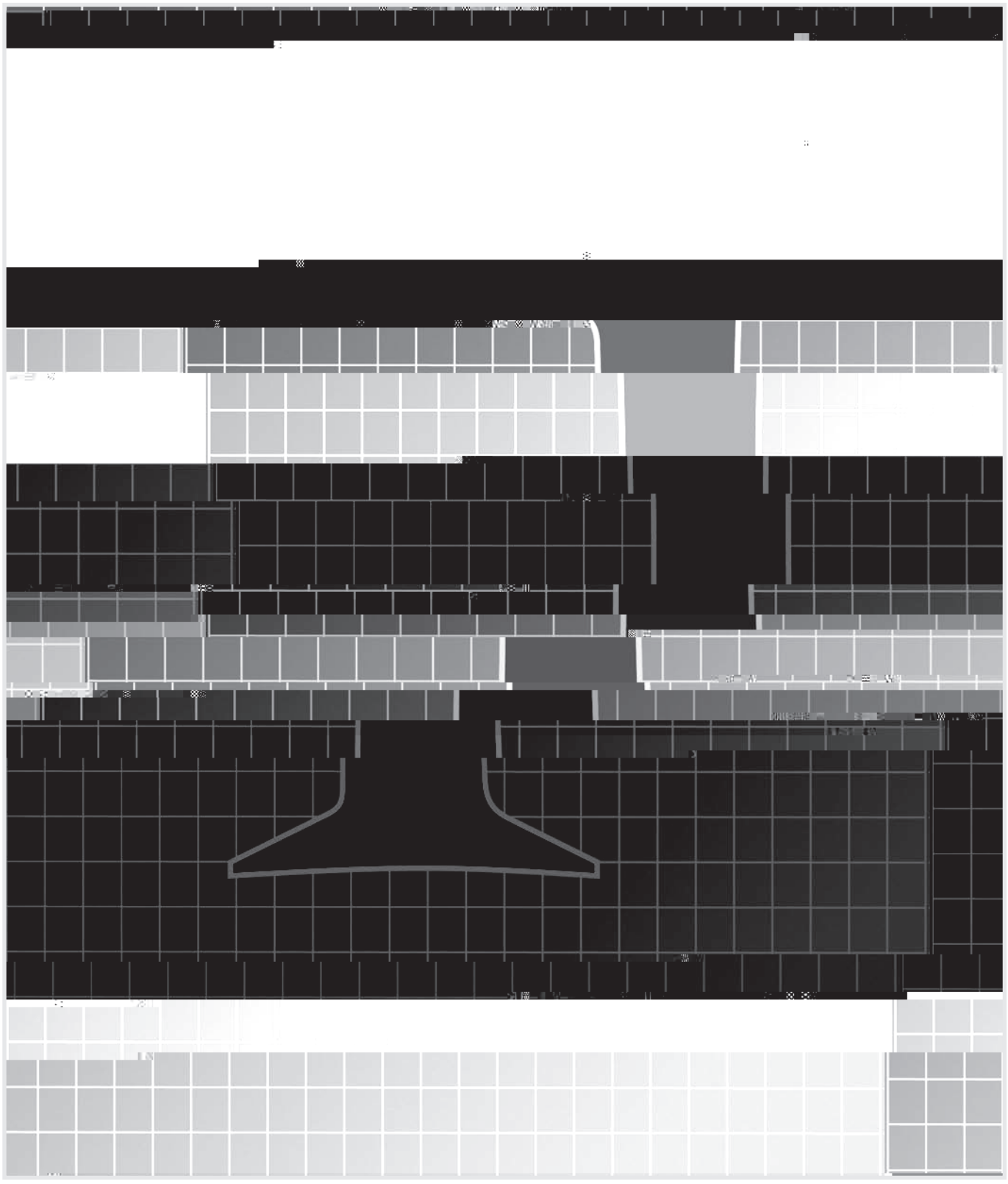
TIMSS & PIRLS International Study Center
Lynch School of Education
Manresa House
Boston College
Chestnut Hill, MA 02467
United States

tel: +1-617-552-1600
fax: +1-617-552-1203

E-mail: timss@bc.edu

URL: timss.bc.edu

Contents



Science Foundation and the US National Center for Education Statistics,

expressed a need for comparative information about how students perform in the cognitive domains. To provide enhanced information from TIMSS 2003 and facilitate planning for TIMSS 2007, a number of participating countries supported a developmental project for IEA's TIMSS & PIRLS International Study Center to examine mathematics achievement by cognitive domains. Although focusing on mathematics as the first step, if successful the project was intended also to serve as a roadmap for achieving similar goals in science.

Led by the United States, with funding also provided by Chinese Taipei, Cyprus, New Zealand, Norway, Ontario, Quebec, Singapore, and Sweden, the developmental project involved several major activities. Prior to preparing this report of the results of the development study, IEA's TIMSS & PIRLS ISC first convened an international meeting of experts in mathematics and mathematics education to confirm the mapping of TIMSS 2003 mathematics items to cognitive domains. Then, IEA's TIMSS & PIRLS ISC conducted the various phases of the analytic work necessary to create the cognitive domain scale scores.

Mapping the TIMSS 2003 Mathematics Items to Cognitive Domains

The developmental project began with a special meeting of mathematics experts held in February 2005 in Amsterdam, with the purpose of examining the classification of items according to the cognitive domains articulated in the TIMSS 2003 mathematics framework. The 10 participants (see Appendix B) expressed great enthusiasm for the meeting goal – facilitating TIMSS reporting according to cognitive domains. Nevertheless, all members expressed reservations about using the cognitive domains as they stood.

In developing the TIMSS 2003 Assessment Framework for Mathematics, there were no plan[illegible]

accomplished at the special Mathematics Cognitive Domains meeting and worked toward refining the classifications and their descriptions to better reflect the essence of the three cognitive domains. This resulted in an excellent foundation for scaling the TIMSS 2003 achievement data by cognitive domains.

Also, IEA's TIMSS & PIRLS ISC examined the distribution of the items within the three cognitive domains by item type, content domain, and difficulty to ensure that there was sufficient coverage of each of the newly defined domains. As described in Appendix B (and summarized in Exhibit B.1), there was a substantial number of items in each domain: 65 in knowing, 93 in applying, and 36 in reasoning at eighth grade; and 58 in knowing, 63 in applying, and 38 in reasoning at fourth grade. Within each domain, there was a good spread of item type (constructed-response or multiple-choice) at both grades, although as might be expected, relatively more of the knowing items were multiple choice and relatively more reasoning items constructed response. There also was a good spread of items across content domains within each of the three cognitive domains, although there was some unevenness in some areas. For example, it would have been preferable to have a higher proportion of number items in the reasoning domain at the eighth grade, and a higher proportion of patterns and relationship items in the knowing domain and measurement items in the reasoning domain at fourth grade. For TIMSS 2007, an effort

further refinements for TIMSS 2007 as published in the *TIMSS 2007 Assessment Frameworks* (Mullis, Martin, Ruddock, O'Sullivan, Arora, and Erberber, 2005).

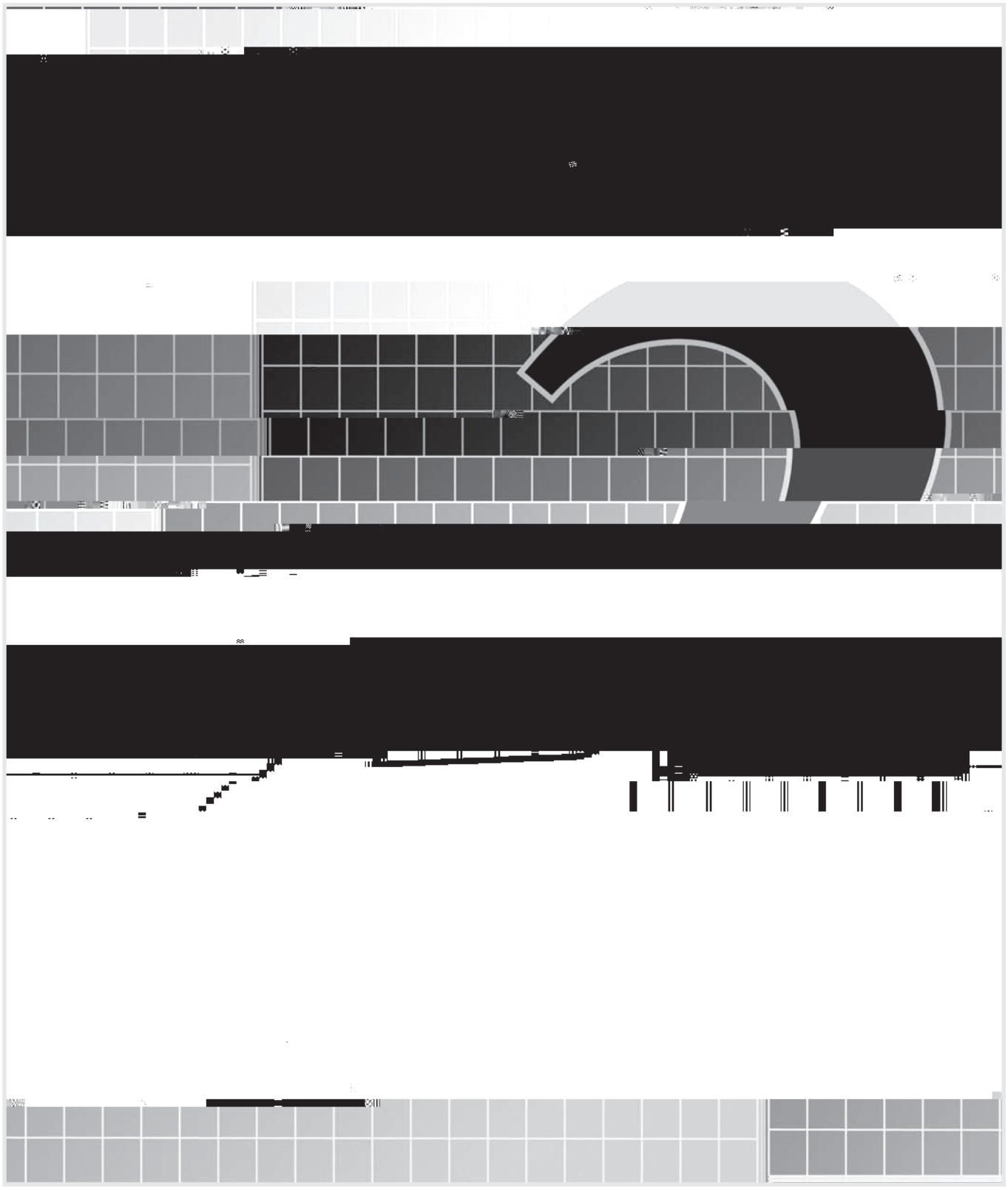
The Scaling Methodology

The methodology used to create the mathematics cognitive domain scales was identical to that used to report mathematics achievement results and achievement in the mathematics content domains in

Summary of Overall Mathematics Achievement Nationally and by Gender for the TIMSS 2003 Countries

To provide a context for considering mathematics achievement at the

to grade. Thus, even though TIMSS devoted considerable effort to maximizing comparability across the grades tested there was some variation. Most notably, in the eighth-grade population, students in Norway, most of Slovenia, and parts of the Russian Federation had fewer years of formal schooling than their counterparts in other countries, while those in England, Scotland, New Zealand, and parts of Australia had more years of schooling. In the fourth-grade population, some students in Slovenia and parts of the Russian Federation had only three years of formal schooling, and students in England and Scotland as well as some in Australia and New Zealand had five years. And, equivalence of chronological a ochr a y ثم a l Scd0



Chapter 2

Mathematics Achievement in the Cognitive Domains at the Fourth and Eighth Grades

This chapter of the report presents the TIMSS 2003 mathematics achievement results for each of the three cognitive domains. Following the presentation of the results, for each domain in turn – knowing, applying, and reasoning – there is an overview of performance across domains.

Knowing Facts, Procedures, and Concepts

The first page of Exhibit 2.1 presents the distribution of students'

At the fourth grade, the difference was also large between the highest-performing country Singapore (626) and the lowest-performing country Tunisia (338). Thirteen countries and the three benchmarking entities performed above the international average and eight countries performed below the international average. The four countries performing about at the international average were Australia, Moldova, Cyprus, and New Zealand.

For both the eighth and fourth grades, Exhibit 2.1 illustrates the broad range of achievement both within and across the countries assessed. It provides a graphical representation of student performance within each country. The bar graph for each country shows the 5th, 25th, 75th, and 95th percentiles[1] as well as the 95% confidence for the mean. Each percentile point indicates the percentage of students below that point on the scale. For most TIMSS 2003 participants at the eighth grade, there was an enormous range within each country between the highest and lowest scores, often as much as 400 scale-score points. This range was as large or larger than the difference in mean achievement between the highest and lowest performing country. For the eighth grade knowing scale, the range for most students in the higher-achiev

Exhibit 2.2: Multiple Comparisons of Average Mathematics Achievement for Knowing Cognitive Domain

MATHEMATICS Grade 8

Countries	- FI=9! 4=G" F>	5ME?9GFI=	* FE? - FE?! 5#4	%@AE=J= 69AG=A	,9G9E	'JKFEA9	$=C?ALD (C=DAJ@	* LE?9I0	0=K@=I9E<J	4LJJA9E (=<=I9KAFE	.9KMA9	5CFM9B 4=GL:CA;	.AK@L9EA9	7EAK=< 5K9K=J	/ 9C9OJA9	+JI9=C	5CFM=EA9	#LJKI9CA9	5=I:A9	'E?C9E<	5N=<=E	$LC?9IA9	4FD9EA9	0=N 8=9C9E<	+K9C0	5,FK9E<	#ID=EA9	/ FC<FM9! 4=G" F>	%OGILJ	OFIN90
- FI=9! 4=G" F>					▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
5ME?9GFI=					▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
* FE? - FE?! 5#4					▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
%@AE=J= 69AG=A					▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
,9G9E	▼	▼	▼	▼		▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
'JKFEA9	▼	▼	▼	▼	▼				▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
$=C?ALD (C=DAJ@	▼	▼	▼	▼	▼				▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
* LE?9I0	▼	▼	▼	▼	▼				▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
0=K@=I9E<J	▼	▼	▼	▼	▼	▼	▼	▼					▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
4LJJA9E (=<=I9KAFE	▼	▼	▼	▼	▼	▼	▼	▼					▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
.9KMA9	▼	▼	▼	▼	▼	▼	▼	▼					▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
5CFM9B 4=GL:CA;	▼	▼	▼	▼	▼	▼	▼	▼						▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
.AK@L9EA9	▼	▼	▼	▼	▼	▼	▼	▼	▼						▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
7EAK=< 5K9K=J	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼				▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
/ 9C9OJA9	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼						▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
+JI9=C	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼					▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
5CFM=EA9	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼					▲	▲	▲	▲	▲	▲	▲	▲	▲	▲	▲
#LJKI9CA9	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼						▲				▲	▲	▲	▲	▲	▲
5=I:A9	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼					▲	▲			▲	▲	▲	▲	▲	▲
'E?C9E<	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼	▼										▲	▲	▲
5N=<=E	▼	▼	▼	▼	▼	▼																							

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

Exhibit 2.2: Multiple Comparisons of Average Mathematics Achievement for Knowing Cognitive Domain

MATHEMATICS Grade 4

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

Exhibit 2.3: Distribution of Mathematics Achievement for Applying Cognitive Domain

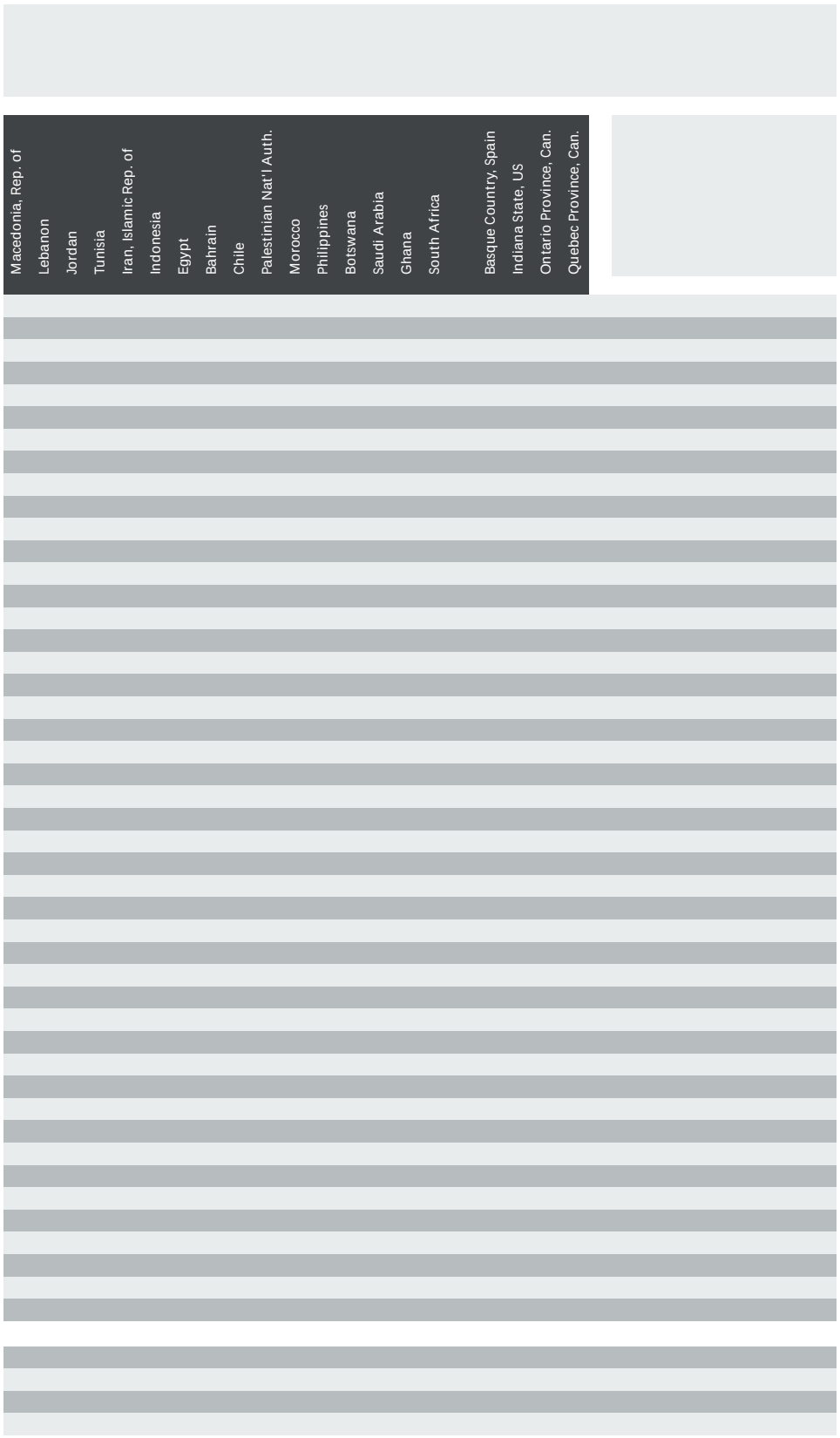
MATHEMATICS Grade 4

Countries	Years of Schooling*	Average Age	Mathematics Achievement Distribution		Human Development Index**
Singapore	4	10.3		595 (5.9) ▲	0.884
† Hong Kong, SAR	4	10.2		577 (3.3) ▲	0.889
Japan	4	10.4		566 (2.1) ▲	0.932
Chinese Taipei	4	10.2		561 (1.9) ▲	–
Belgium (Flemish)	4	10.0		546 (2.1) ▲	0.937
Latvia	4	11.1		545 (3.3) ▲	0.811
Russian Federation	3 or 4	10.6		542 (4.7) ▲	0.779
† Lithuania	4	10.9		542 (2.9) ▲	0.824
† Netherlands	4	10.2		541 (2.6) ▲	0.938
Hungary	4	10.5		530 (3.4) ▲	0.837
† England	5	10.3		526 (4.1) ▲	0.930
Cyprus	4	9.9		510 (2.8) ▲	0.891
Moldova, Rep. of	4	11.0		507 (4.8) ▲	0.700
† United States	4	10.2		505 (2.6) ▲	0.937
International Avg.	4	10.3		495 (0.7)	–
Italy	4	9.8		494 (3.6)	0.916
† Australia	4 or 5	9.9		490 (3.8)	0.939
† Scotland	5	9.7		487 (3.5) ▼	0.930
New Zealand	4.5 - 5.5	10.0		486 (2.3) ▼	0.917
Slovenia	3 or 4	9.8		477 (2.8) ▼	0.881
Armenia	4	10.9		462 (3.2) ▼	0.729
ø Norway	4	9.8		446 (2.2) ▼	0.944
Iran, Islamic Rep. of	4	10.4		391 (3.8) ▼	0.719
Philippines	4	10.8		364 (7.5) ▼	0.751
Morocco	4	11.0		349 (4.5) ▼	0.606
Tunisia	4	10.4		348 (4.6) ▼	0.740
Benchmarking Participants					
Indiana State, US	4	10.5	¾ · (4.5)	523(0.93	

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

Exhibit 2.4: Multiple Comparisons of Average Mathematics Achievement for Applying Cognitive Domain

MATHEMATICS Grade 8



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

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

Exhibit 2.6: Multiple Comparisons of Average Mathematics Achievement for Reasoning Cognitive Domain

MATHEMATICS Grade 8

Countries	Singapore	Korea, Rep. of	Chinese Taipei	Japan	Hong Kong, SAR	Netherlands	Belgium (Flemish)	Hungary	Estonia	Australia	Scotland

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

Exhibit 2.6: Multiple Comparisons of Average Mathematics Achievement for Reasoning Cognitive Domain

MATHEMATICS Grade 4

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

country's average mathematics achievement in the reasoning domain compares to achievement in the other participating countries.

At the eighth grade, average achievement in the reasoning domain ranged from 583 in Singapore to 287 in South Africa. Twenty-four countries and the four benchmarking participants performed significantly above the international average, three countries (Bulgaria, Armenia, and Serbia) performed comparably to the international average, and 19 countries performed significantly below the average.

At the eighth grade, looking at both Exhibits 2.5 and 2.6, it can be seen that the rank ordering of significant differences in achievement is rather complicated for the reasoning domain. Singapore and the Republic of Korea had the highest average achievement in the reasoning domain, nearly identical (583 and 582), but Singapore had a larger standard error (3.5 to 1.7). Thus, the Republic of Korea had significantly higher achievement than every participating country except Singapore and Chinese Taipei whereas Singapore (with the larger standard error) had higher average achievement than every participating country except the Republic of Korea, Chinese Taipei, and Japan. Chinese Taipei and Japan had the same average score (576) followed by Hong Kong SAR. Chinese Taipei (also with a relatively larger standard error of 4.2) did not perform statistically differently than the other three Asian countries, whereas a difference was found between the Republic of Korea and Japan due to their small standard errors. Hong Kong SAR was outperformed only by Singapore and the Republic of Korea. The Netherlands and Belgium (Flemish) only were outperformed by the five top-scoring Asian countries.

At the fourth grade, performance ranged from 574 for Singapore to 340 for Tunisia. Fifteen countries and the three benchmarking participants performed significantly above the international average, three countries (Italy, Scotland, and Moldova) performed essentially at the international average, and seven countries performed significantly below the international average. Singapore had the highest achievement, outperforming all countries except Hong Kong SAR and

Chinese Taipei. Hong Kong, Chinese Taipei and Japan had similar achievement followed by Belgium (Flemish), England, and the Netherlands (all with similar average achievement and only outperformed by the four highest-achieving Asian countries).

Overview Across Domains

At both the eighth and fourth grades, the countries with the highest achievement in each of the three cognitive domains also tended to be the highest-scoring countries (though not always in the same rank order) on the overall mathematics assessment. At the eighth grade (see Exhibit 1.1), the four countries with the highest overall mathematics

nH ƌ ƕnk

- In reasoning, Singapore and the Republic of Korea performed very similarly followed by Chinese Taipei and Japan and then Hong Kong SAR (see Exhibits 2.5 and 2.6). The Netherlands, Belgium (Flemish), and the Canadian province of Quebec also had relatively high achievement, only being outperformed by the five Asian countries.

At the fourth grade, Singapore was the highest-performing country in overall mathematics followed by Hong Kong SAR, and then by Japan and Chinese Taipei who performed similarly (see Exhibit 1.1). Belgium (Flemish) had higher achievement than all countries except these four Asian countries.

- In knowing, the pattern at the fourth grade was the same as for overall mathematics (see Exhibits 2.1 and 2.2). The four Asian countries had the best achievement (Singapore followed by Hong Kong SAR, and then by Chinese Taipei and Japan) with Belgium (Flemish) having higher achievement than all countries except the four best-achieving Asian countries.
- In applying, the pattern for the four high-achieving Asian countries was the same as for overall mathematics (see Exhibits 2.3 and 2.4). However, Belgium (Flemish), Latvia, the Russian Federation, Lithuania, and the Netherlands all followed, performing similarly to each other with lower achievement than the four Asian countries, but higher achievement than the rest of the participating countries.
- In reasoning, Singapore, Hong Kong SAR, and Chinese Taipei had the highest achievement (see Exhibits 2.5 and 2.6). Japan had achievement similar to Hong Kong SAR and Chinese Taipei, but was outperformed by Singapore. Belgium (Flemish), England, and the Netherlands had achievement equal to or higher than all participants except the four top-achieving Asian countries.

Just as countries with high achievement on the mathematics assessment as a whole had high achievement in the three cognitive dti

(e.g., South Africa, Ghana, and Saudi Arabia at the eighth grade and the Philippines, Morocco, and Tunisia at the fourth grade) also

Gender Differences in the Applying Cognitive Domain

Exhibit 3.2 shows achievement differences between girls and boys for the applying domain for the eighth and fourth grades, (on the first and second pages, respectively). For the applying domain at the eighth grade, boys had significantly higher achievement in more countries than girls. Girls had significantly higher achievement than boys in seven countries, and boys had significantly higher achievement than girls in 13 countries and two benchmarking participants (the US state of Indiana and the Canadian province of Quebec).

Fourth grade had a corresponding pattern for the applying domain, with boys having significantly higher achievement in more countries than girls. Girls had higher achievement in the applying domain in four countries whereas boys had higher achievement in seven countries and the two Canadian provinces. Also, internationally, on average, there was a small but significant difference favoring boys.

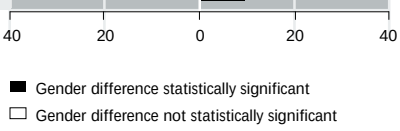
							Girls Scored Higher	Boys Scored Higher
	Norway	50 (0.8)	469 (2.8)	50 (0.8)	468 (3.4)	0 (3.2)		
	Slovenia	50 (0.9)	491 (3.0)	50 (0.9)	491 (2.8)	0 (3.6)		
	Russian Federation	49 (1.2)	503 (3.8)	51 (1.2)	503 (4.1)	0 (2.6)		
†	Hong Kong, SAR	50 (2.4)	584 (3.7)	50 (2.4)	584 (4.5)	1 (5.1)		
	Romania	52 (0.9)	475 (5.4)	48 (0.9)	474 (5.3)	1 (3.9)		
1	Indonesia	50 (0.7)	408 (5.0)	50 (0.7)	409 (5.3)	1 (3.3)		
	New Zealand	52 (1.7)	496 (4.7)	48 (1.7)	497 (7.2)	1 (5.9)		
	Bulgaria	48 (1.3)	471 (6.0)	52 (1.3)	472 (4.9)	1 (5.5)		
	International Avg.	50 (0.2)	466 (0.6)	50 (0.2)	467 (0.6)	1 (0.6)		
	Sweden	51 (0.9)	504 (3.2)	49 (0.9)	506 (2.8)	1 (2.2)		
	Palestinian Nat'l Auth.	55 (2.4)	389 (4.1)	45 (2.4)	388 (4.6)	1 (5.8)		
	Latvia	49 (0.8)	505 (3.5)	51 (0.8)	504 (4.1)	2 (3.4)		
1	Lithuania	50 (0.9)	499 (3.2)	50 (0.9)	497 (3.3)	2 (2.9)		
	Botswana	51 (0.7)	370 (3.0)	49 (0.7)	368 (2.9)	2 (2.4)		
	Japan	49 (1.2)	563 (4.4)	51 (1.2)	565 (3.6)	2 (6.7)		
†	Scotland	50 (1.3)	506 (4.8)	50 (1.3)	504 (3.8)	3 (3.8)		
1	Serbia	49 (0.8)	468 (3.5)	51 (0.8)	466 (3.1)	3 (2.9)		
	South Africa	51 (0.9)	267 (5.9)	49 (0.9)	271 (6.5)	3 (6.1)		
	Estonia	50 (1.0)	531 (3.3)	50 (1.0)	526 (3.2)	4 (2.9)		
	Chinese Taipei	48 (1.0)	584 (5.1)	52 (1.0)	580 (5.1)	4 (4.2)		
	Egypt	46 (2.7)	401 (4.3)	54 (2.7)	406 (4.9)	5 (6.3)		
2	Macedonia, Rep. of	49 (0.9)	431 (4.2)	51 (0.9)	426 (4.3)	6 (3.9)		
	Slovak Republic	48 (1.3)	499 (4.0)	52 (1.3)	505 (4.3)	6 (3.6)		
◆	Korea, Rep. of	48 (2.8)	581 (2.9)	52 (2.8)	587 (2.3)	6 (2.9)		
	Malaysia	50 (1.8)	515 (5.1)	50 (1.8)	508 (4.8)	7 (4.6)		
	Iran, Islamic Rep. of	40 (4.1)	420 (4.6)	60 (4.1)	413 (4.1)	7 (7.2)		
	Italy	50 (0.9)	479 (3.0)	50 (0.9)	488 (4.0)	8 (3.0)		
	Armenia	53 (0.7)	482 (3.5)	47 (0.7)	473 (3.5)	8 (3.6)		
‡	United States	52 (0.7)	497 (3.5)	48 (0.7)	506 (3.5)	9 (2.1)		
	Moldova, Rep. of	51 (0.8)	462 (4.0)	49 (0.8)	453 (4.5)	9 (3.3)		
	Lebanon	57 (1.8)	422 (3.7)	43 (1.8)	432 (4.2)	10 (4.0)		
†	Netherlands	49 (1.2)	538 (4.0)	51 (1.2)	548 (4.3)	10 (3.8)		
	Philippines	58 (0.9)	383 (4.8)	42 (0.9)	373 (5.5)	10 (3.5)		
2	Israel	52 (1.6)	490 (3.7)	48 (1.6)	500 (4.6)	10 (4.2)		
	Singapore	49 (0.8)	617 (3.6)	51 (0.8)	606 (4.1)	11 (3.1)		
	Hungary	50 (1.0)	517 (3.8)	50 (1.0)	529 (4.0)	11 (3.5)		
	Saudi Arabia	43 (2.3)	332 (6.1)	57 (2.3)	344 (4.5)	12 (7.9)		
	Ghana	45 (0.9)	286 (4.9)	55 (0.9)	299 (4.8)	13 (5.2)		
	Australia	51 (2.2)	501 (6.1)	49 (2.2)	516 (6.0)	15 (7.5)		
	Belgium (Flemish)	54 (2.1)	529 (3.3)	46 (2.1)	544 (3.7)	15 (4.6)		
	Cyprus	49 (0.6)	465 (1.9)	51 (0.6)	450 (2.5)	16 (3.1)		
1 ‡	Morocco	50 (1.8)	377 (3.4)	50 (1.8)	393 (3.3)	16 (3.4)		
	Chile	48 (1.6)	382 (3.6)	52 (1.6)	399 (4.2)	18 (4.6)		
	Tunisia	53 (0.7)	407 (2.6)	47 (0.7)	433 (2.4)	26 (2.1)		
	Jordan	49 (1.7)	436 (4.9)	51 (1.7)	409 (5.8)	27 (6.9)		
	Bahrain	50 (0.4)	411 (2.3)	50 (0.4)	384 (2.3)	27 (3.2)		
‡	England	50 (2.4)	503 (5.4)	50 (2.4)	504 (6.0)	1 (2.2)		

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

Exhibit 3.2: Average Mathematics Achievement by Gender for Applying Cognitive Domain

MATHEMATICS Grade 4

Countries	Girls		Boys		Difference (Absolute Value)	Gender Difference	
	Percent of Students	Average Scale Score	Percent of Students	Average Scale Score		Girls Scored Higher	Boys Scored Higher
Iran, Islamic Rep. of	39 (4.2)	391 (6.1)	61 (4.2)	391 (4.8)	0 (7.9)		
† Hong Kong, SAR	47 (1.1)	576 (3.5)	53 (1.1)	577 (3.5)	0 (2.6)		
Chinese Taipei	48 (0.5)	561 (2.0)	52 (0.5)	562 (2.2)	1 (2.0)		
Latvia	49 (0.9)	545 (3.4)	51 (0.9)	546 (4.0)	1 (3.4)		
New Zealand	50 (1.1)	485 (3.0)	50 (1.1)	486 (2.5)	1 (3.2)		
Hungary	50 (0.9)	530 (4.0)	50 (0.9)	531 (3.7)	2 (3.7)		
International Avg.	49 (0.2)	494 (0.8)	51 (0.2)	497 (0.8)	2 (0.8)		
† England	50 (0.9)	524 (4.1)	50 (0.9)	528 (4.9)	4 (4.0)		
Belgium (Flemish)	50 (1.0)	544 (2.5)	50 (1.0)	548 (2.7)	4 (3.0)		
† Lithuania	49 (0.9)	541 (3.7)	51 (0.9)	545 (3.7)	4 (3.7)		
† Australia	50 (1.0)	487 (4.3)	50 (1.0)	492 (4.4)	5 (3.9)		
Japan	49 (0.6)	563 (2.6)	51 (0.6)	569 (2.3)	5 (2.5)		
Tunisia	48 (0.9)	351 (5.1)	52 (0.9)	346 (4.7)	6 (3.6)		
Russian Federation	50 (0.7)	539 (5.0)	50 (0.7)	545 (5.1)	6 (3.8)		
Norway	50 (0.8)	443 (2.8)	50 (0.8)	449 (3.0)	6 (3.8)		
Armenia	49 (0.8)	465 (3.2)	51 (0.8)	459 (3.7)	6 (2.8)		
† Netherlands	49 (1.1)	538 (2.8)	51 (1.1)	545 (3.2)	7 (3.1)		
Morocco	49 (1.1)	345 (5.6)	51 (1.1)	352 (4.4)	7 (4.5)		
Slovenia	48 (1.1)	474 (3.1)	52 (1.1)	481 (3.8)	7 (4.2)		
Italy	48 (0.8)	489 (4.3)	52 (0.8)	498 (3.5)	8 (3.0)		
Singapore	49 (1.4)	599 (5.8)	51 (1.4)	590 (6.6)	9 (4.1)		
Moldova, Rep. of	50 (0.8)	511 (5.3)	50 (0.8)	502 (5.0)	9 (3.6)		
† United States	50 (0.5)	501 (2.8)	50 (0.5)	510 (2.9)	9 (2.1)		
† Scotland	51 (1.0)	482 (3.6)	49 (1.0)	492 (4.6)	11 (4.3)		
Cyprus	49 (0.7)	504 (3.1)	51 (0.7)	516 (2.9)	12 (2.4)		
Philippines	51 (1.0)	370 (8.8)	49 (1.0)	357 (6.7)	13 (4.8)		
Benchmarking Participants							
Indiana State, US	52 (1.1)	521 (3.4)	48 (1.1)	525 (4.2)	4 (3.8)		
Ontario Province, Can.	48 (1.1)	491 (3.8)	52 (1.1)	505 (5.8)	14 (4.4)		
Quebec Province, Can.	50 (0.9)	493 (3.2)	50 (0.9)	503 (2.9)	9 (2.8)		



■ Gender difference statistically significant
□ Gender difference not statistically significant

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

¢ Kcr esgbcjg lcq dmp q_ knjc n_prqaqn_rqm l p_rcq m ljw _drcp pcnj_ack clr qafmmjq ucpc g lajsbcb &qcc Cvf g`gr A,0',

/ L_rqm l_j Bcqgpcb Nmnsj_rqm l bmcq lmr amtcp _jj md &lrcp l_rqm l_j Bcqgpcb Nmnsj_rqm l &qcc Cvf g`gr A,/',

& ' Qr_lb_pb cppmpq _nnc_p g l n_pc lrfcqcq, @ca_sqc pcqsjrq _pc pms lbcb rm rfc lc_pcqr u fmjc lsk `cp* qm kc rmr_jq k_w _nnc_p g lam lqgqrc lr,

Gender Differences in the Reasoning Cognitive Domain

Exhibit 3.3 shows gender achievement differences in the reasoning domain at the eighth grade (first page) and fourth grade (second page). On average, across all countries, eighth-grade girls had significantly higher achievement than boys in the reasoning domain. In this domain, girls had significantly higher achievement than boys in 17 countries and the Basque Country, Spain whereas boys had higher achievement in only two countries (Morocco and Tunisia).

At the fourth grade this pattern was similar, but far less pronounced. There was essentially no difference in achievement internationally between fourth-grade boys and girls in the reasoning domain. However, girls had higher achievement than boys in three countries whereas boys did not outperform girls in any country or benchmarking entity.

Chapter 4

Country by Country Profiles of Achievement in the Mathematics Cognitive Domains

To highlight relative strengths and weaknesses within each country, this chapter describes in which mathematics cognitive areas each country is relatively strong or weak. Regardless of international standing, the profiles of achievement within country reveal that many countries performed relatively better or worse in one or more cognitive domains than they did overall.

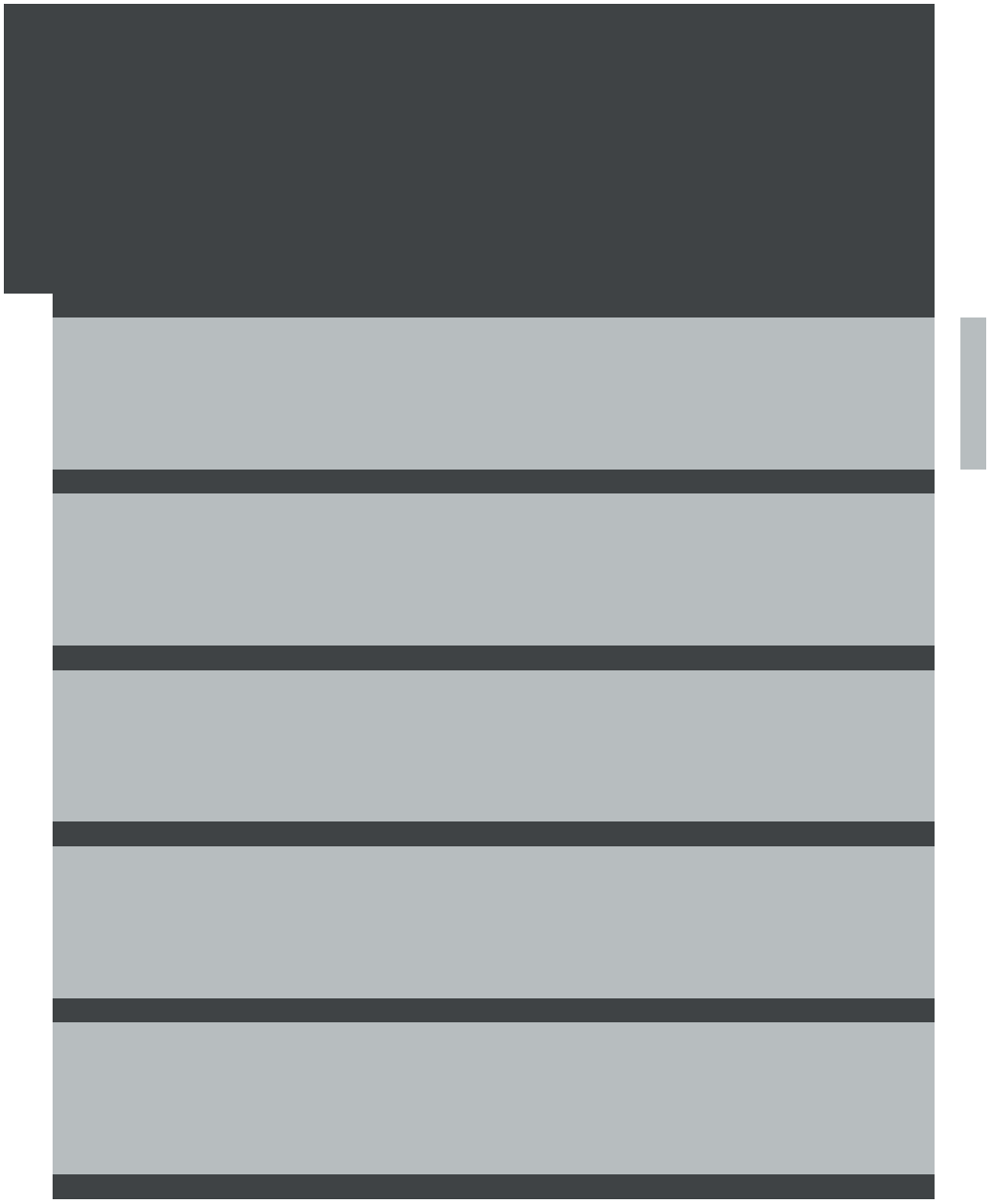
Differences in relative performance may be related to one or more of a number of factors, such as emphases in intended curriculum or widely used textbooks, differences in instruction and curriculum implementation, and differences in the match between instruction and the types of items contained in TIMSS 2003.

Profiles of Achievement

For each country, Exhibit 4.1 displays the difference between average performance in each content area and the country's average performance overall. The first three pages of Exhibit 4.1 show the results for eighth grade and the next two pages show the results for the

fourth grade. For each country, the average of the cognitive domain scores has been set to zero, so that above average or below average performance can be highlighted for each of the three domains. Relatively better achievement in a cognitive domain is shown when the circle and the lines indicating its confidence interval are completely above and not touching zero on the scale, and relatively worse achievement by a circle and its confidence interval lines completely below "0."

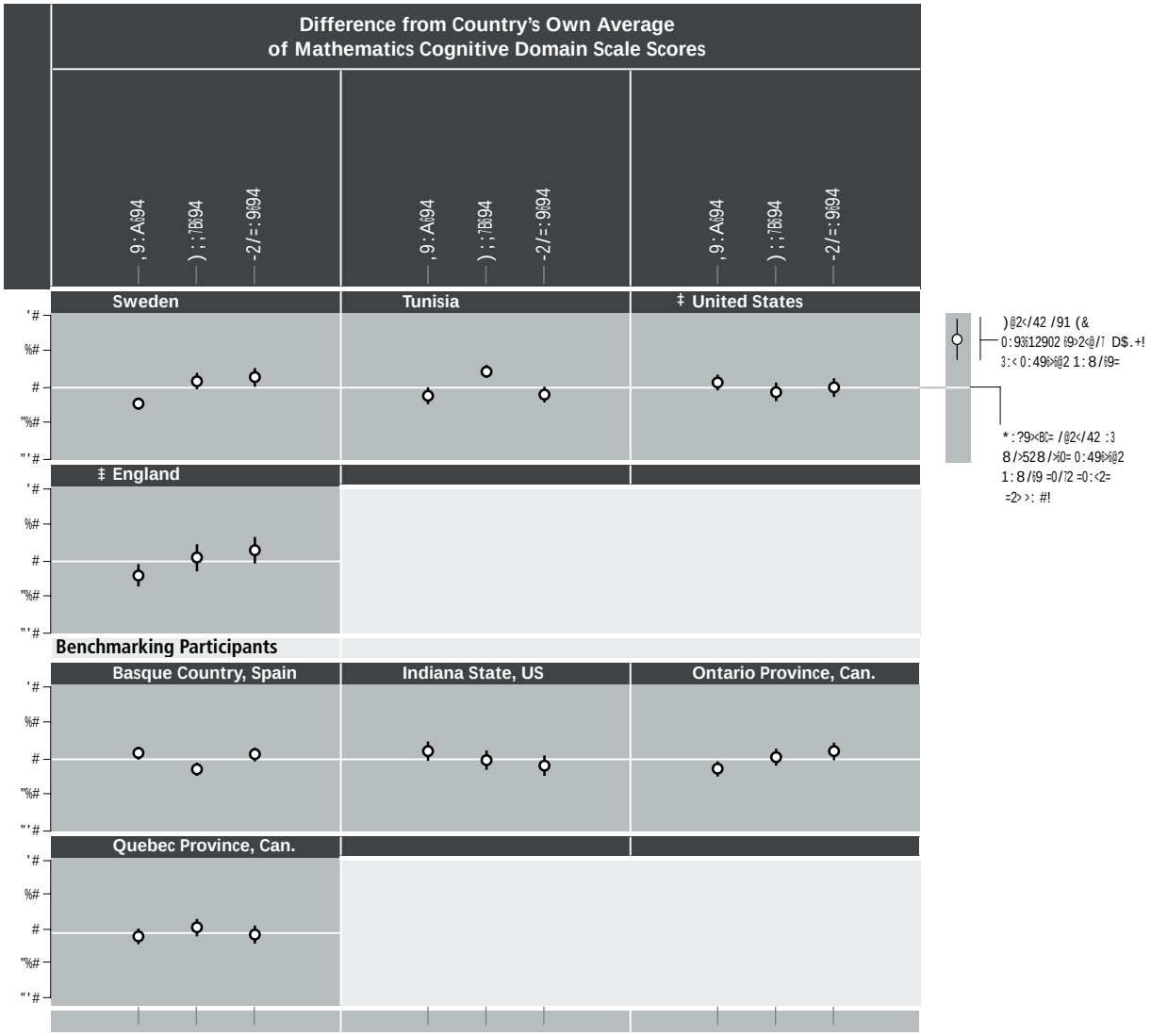
The profiles of relative performance reveal interesting differences among countries. Most countries show the profile of performing



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

Exhibit 4.1: Profiles of Within-Country Relative Performance in Mathematics Cognitive Domains

MATHEMATICS Grade 8



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

£ Lc_pjw q_rqqĝcb esgbcjgl cq dmp q_knjc n_prqaqn_rqml p_rcq mljw _drcp pcnj_ackclr qafmmjq ucpc gl ajsbcb &qcc Cvfg`gr A,0',

Ë Bgb lmr q_rqqdw esgbcjgl cq dmp q_knjc n_prqaqn_rqml p_rcq &qcc Cvfg`gr A,0',

fourth grade, countries with a relative strength in the knowing domain were Belgium (Flemish), Italy, Singapore, the United States, and the US state of Indiana. Comparatively more countries at the fourth grade had a relative weakness in the knowing domain, including Cyprus, Latvia, Lithuania, the Netherlands, Norway, the Russian Federation, and Slovenia.

Relative Strengths and Weaknesses in the Applying Domain

At the eighth grade, there were fewer countries with differences between overall mathematics achievement and achievement in the applying domain than there were with such differences in the knowing domain. Countries with a relative strength in the applying domain at the eighth grade included Ghana, Singapore, and Tunisia. Those with a relative weakness in the applying domain included Bahrain, Macedonia, and Serbia.

At the fourth grade, Armenia, Latvia, Lithuania, and the Russian Federation had applying as a particular strength. Compared to performance in overall mathematics, applying was a relative weakness in Australia, New Zealand, Norway, the United States, the US state of Indiana, and the two Canadian provinces (a group including three English-speaking countries).

Relative Strengths and Weaknesses in the Reasoning Domain

Countries with the reasoning domain as a particular strength at the eighth grade included Bahrain, Chile, Ghana, Japan, Norway, the Palestinian National Authority, Saudi Arabia, Scotland, South Africa, and Sweden. Countries that performed less well in the reasoning domain than they did in overall mathematics included Armenia, Botswana, Cyprus, Hong Kong, Israel, Lebanon, Lithuania, the Philippines, Romania, the Russian Federation, Serbia, and Singapore.

At the fourth grade, the participants with a relative strength in reasont

the two Canadian provinces. Only two countries, Belgium (Flemish) and Singapore, did relatively less well in reasoning as compared to their overall mathematics performance.

International Achievement Across the Cognitive Domains

At the eighth grade across the TIMSS 2003 participants, the knowing domain had the most differences, with many countries showing either a relative strength or weakness in this area. Fifteen countries performed better in the knowing domain than they did in mathematics overall, and 12 countries and the Canadian province of Ontario performed worse. The applying domain was the cognitive area least likely to feature either relatively strong or relatively weak performance. Only three countries performed better in the applying domain than they did in mathematics overall (Ghana, Singapore, and Tunisia) and only three countries performed worse (Bahrain, Macedonia, and Serbia).

In the reasoning domain at the eighth grade, 10 countries performed relatively better than they did in mathematics overall and 12 countries did less well. The countries making up each of the two groups included those from very different parts of the world geographically and with disparate cultures and mathematics traditions. For example, the countries with a relative strength in the reasoning domain were Bahrain, Chile, Ghana, Japan, Norway, the Palestinian National Authority, Saudi Arabia, Scotland, South Africa, and Sweden.

At the fourth grade, looking across the participating countries, about the same number of differences (strengths or weaknesses) occurred in each of the cognitive domains. However, several more countries showed a relative weakness in the knowing cognitive domain (seven) than had this domain as a relative strength (five). Similarly, more countries had a relative weakness in the applying domain (seven) than had this domain as a relative strength (four). In comparison, more countries showed a relative strength in the reasoning domain (seven) than showed this domain as a relative weakness (two).

References

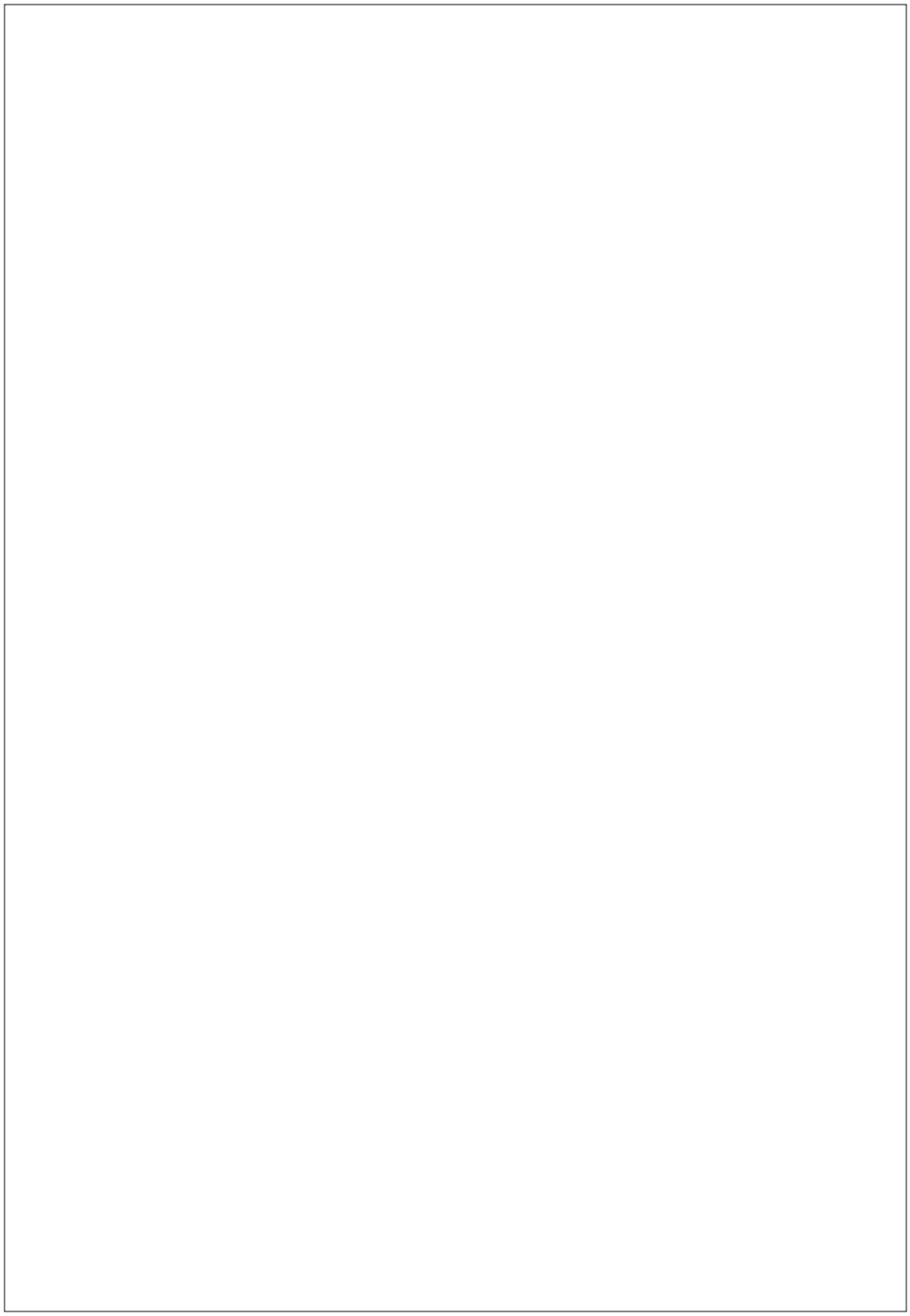
Gonzalez, E.J., Galia, J., & Li, I. (2004). Scaling methods and procedures for the TIMSS 2003 mathematics and science scales. In M.O. Martin, I.V.S. Mullis, & S.J. Chrostowski (Eds.), *TIMSS 2003 technical report*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.

Martin, M.O., Mullis, I.V.S., & Chrostowski, S.J. (2004), *TIMSS 2003 technical report*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.

Mullis, I.V.S., Martin, M.O., Smith, T.A., Garden, R.A., Gregory, K.D., Gonzalez, E.J., Chrostowski, S.J., & O'Connor 00S

Without access to a knowledge base that enables easy recall of the language and basic facts and conventions of number, symbolic representation, and spatial relations, students would find purposeful mathematical thinking impossible. *Facts* encompass the factual knowledge that provides the basic language of mathematics, and the essential mathematical facts and properties that form the foundation for mathematical thought.

Procedures form a bridge between more basic knowledge and the use of mathematics for solving routine problems, especially those encountered by many people in their daily lives. In essence a fluent use of procedures entails recall of sets of actions and how to carry them out. Students need to be efficient and accurate in using a variety of



standard in classroom exercises designed to provide practice in particular methods or techniques. Some of these problems will have been in words that set the problem situation in a quasi-real context. Though they range in difficulty, each of these types of "textbook" problems is expected to be sufficiently familiar to students that they will essentially involve selecting and applying learned procedures.

Problems may be set in real-life situations, or may be concerned with purely mathematical questions involving, for example, numeric or algebraic expressions, functions, equations, geometric figures, or statistical data sets. Therefore, problem solving is included not only in the *applying knowledge and conceptual understanding* domain, with emphasis on the more familiar and routine tasks, but also in the *reasoning* domain.

This cognitive domain covers the following behaviors:

Select	Select an efficient/appropriate operation, method or strategy for solving problems where there is a known algorithm or method of solution. Select simple algebraic expressions which represent straightforward situations (fourth grade). Select the nets of simple geometric figures (fourth grade). Select appropriate algorithms or formulas.
Represent	Display mathematical information and data in diagrams, tables, charts, or graphs, and generate equivalent representations for a given mathematical entity or relationship.

Solve Routine Problems	Solve routine problems (i.e., problems similar to those target students are likely to have encountered in class). For example, use geometric properties to solve problems; compare and match different representations of data (eighth grade) and use data from charts, tables, graphs, and maps to solve routine problems.

Reasoning

Reasoning mathematically involves the capacity for logical, systematic thinking. It includes intuitive and inductive reasoning based on patterns and regularities that can be used to arrive at solutions to non-routine problems. Non-routine problems are problems that are very likely to be unfamiliar to students. They make cognitive demands over and above those needed for solution of routine problems, even when the knowledge and skills requi e n ơ sê v huid lrĦk a e ṗ ge$ a lĠi e

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*

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 4

Number of Items by Item Type and Cognitive Domains

Item Type	Cognitive Domains			Total
	Knowing	Applying	Reasoning	
Constructed Response	&	! '	!#	%(
Multiple Choice	#	"$	#	(
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	!#	" '	" '	&%
Multiple Choice	#%	"(	%	(
Total	**36%**	**39%**	**26%**	**166**

Number of Items by Content Domain and Cognitive Domain

Content Domain	Cognitive Domains			Total
	Knowing	Applying	Reasoning	
Number	!$	(	(	%"
Patterns and Relationships	!	"	'	!"
Measurement		'	#	"!
Geometry	&	%		!#
Data	#	&	%	&
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	"&	!(	"#	%'
Patterns and Relationships	'	$#	" '	!#
Measurement	"	$%	"	"!
Geometry	&!	!#	#	!$
Data	!#	#	"$	&
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	%"	$%	"&	$"
Patterns and Relationships	%"	$&	"%	$
Measurement	%$	#&	"(	$!
Geometry	%	$"	#"	$&
Data	%(	$%	$'	%
Total	**53%**	**53%**	**40%**	**54**

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 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.

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

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

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

	Overall	Knowing	Applying	Reasoning
, N J DK H@ ÀŒ@ #-	0.99 ñ Kñ à ' 4H6ñ	qt¿ñN H2D ` P À à	0.99	#"+)
, QOPN@IH@	0.97	0.98	0.98	#"+$
- DIF HQ J 0ID J HOG	0.99	0.83	0.76	#"+$
. GHKD0D =@HMDH	0.94	0.88	0.89	#")+
4 31 43				
97 0.83 99 1 ø(HOÐ ` P À à				0.98

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

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

MATHEMATICS Grade 8

	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	A Ţ Ã ¾(F	(FleminAR	((FC

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

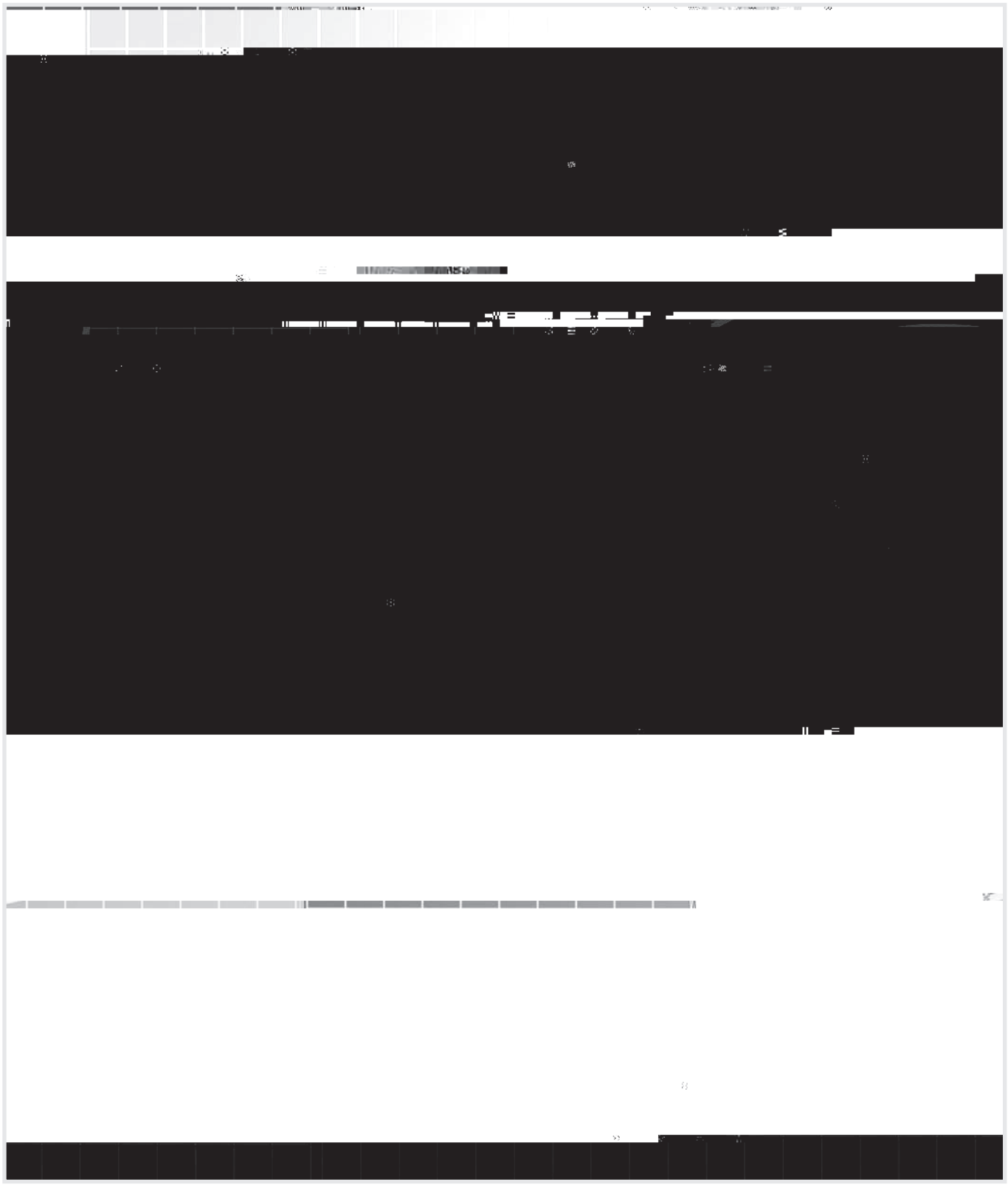
SOURCE: IEA's Trends in International Mathematics and Science Study (TIMSS) 2003

Exhibit B.4: Correlations of Mathematics Cognitive Domains

MATHEMATICS Grade 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



Appendix C

@lsbo^db lc QF J PP /--0 Q^odbq Mlmri^qflkp ^ka M^oqf`fm^qflk O^qbp

Notes on Coverage				
Armenia	100%	2.9%	0.0%	2.9%
Australia	100%	0.4%	0.9%	1.3%
Bahrain	100%	0.0%	0.0%	0.0%
Belgium (Flemish)	100%	3.1%	0.1%	3.2%
Botswana	100%	0.8%	2.2%	3.0%
Bulgaria	100%	0.5%	0.0%	0.5%
Chile	100%	1.6%	0.7%	2.2%
Chinese Taipei	100%	0.2%	4.6%	4.8%
Cyprus	100%	1.1%	1.5%	2.5%
Egypt	100%	3.4%	0. h	

cv z Àaö– – ` cC˜` ` ° ` ð p 0 ° 101 3.4% 0. h

100%

î E.F(A. } } ضǚ Î m̦ m î f

ĥ .B(B.G(

Exhibit C.2: Participation Rates (Weighted)

MATHEMATICS Grade 4

Countries	School Participation		Class Participation	Student Participation	Overall Participation	
	Before Replacement	After Replacement			Before Replacement	After Replacement
Armenia	99%	99%	100%	91%	90%	90%
Australia	78%	90%	100%	94%	73%	85%
Belgium (Flemish)	89%	99%	100%	98%	87%	97%
Chinese Taipei	100%	100%	100%	99%	99%	99%
Cyprus	100%	100%	100%	97%	97%	97%
England	54%	82%	100%	93%	50%	76%
Hong Kong, SAR	77%	88%	99%	95%	73%	83%
Hungary	98%	99%	100%	94%	92%	93%
Iran, Islamic Rep. of	100%	100%	100%	98%	98%	98%
Italy	97%	100%	100%	97%	93%	97%
Japan	100%	100%	100%	97%	97%	97%
Latvia	91%	94%	100%	94%	85%	88%
Lithuania	92%	96%	99%	92%	84%	87%
Moldova, Rep. of	97%	100%	100%	97%	94%	97%
Morocco	87%	87%	100%	93%	81%	81%
Netherlands	52%	87%	100%	96%	50%	84%
New Zealand	87%	98%	100%	95%	82%	93%
Norway	89%	93%	100%	95%	85%	88%
Philippines	78%	85%	100%	95%	75%	81%
Russian Federation	99%	100%	100%	97%	96%	97%
Scotland	64%	83%	100%	92%	59%	77%
Singapore	100%	100%	100%	98%	98%	98%
Slovenia	95%	99%	100%	92%	87%	91%
Tunisia	100%	100%	100%	99%	99%	99%
United States	70%	82%	99%	95%	66%	78%
Yemen	100%	100%	100%	93%	93%	93%
Benchmarking Participants						
Indiana State, US	100%	100%	100%	98%	98%	98%
Ontario Province, Can.	89%	94%	100%	96%	85%	90%
Quebec Province, Can.	99%	100%	100%	91%	90%	91%

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

Exhibit D.1: Percentiles of Achievement in Knowing Cognitive Domain

MATHEMATICS Grade 4

Countries					
Armenia	296 (7.6)	387 (5.9)	447 (5.4)	507 (5.0)	597 (4.6)
Australia	363 (8.0)	449 (5.7)	505 (4.6)	554 (3.3)	631 (5.7)
Belgium (Flemish)	450 (3.4)	512 (2.3)	558 (2.3)	604 (3.7)	665 (2.9)
Chinese Taipei	443 (4.5)	518 (2.6)	566 (2.2)	614 (3.1)	681 (4.1)
Cyprus	351 (3.9)	443 (4.6)	503 (4.3)	560 (2.4)	641 (5.6)
England	378 (7.7)	472 (4.4)	536 (5.1)	599 (5.0)	687 (4.5)
Hong Kong, SAR	458 (4.5)	528 (3.6)	576 (3.6)	620 (4.0)	683 (3.7)
Hungary	382 (4.2)	464 (3.2)	520 (3.5)	572 (4.8)	644 (5.7)
Iran, Islamic Rep. of	275 (8.3)	351 (3.2)	403 (3.3)	457 (5.0)	533 (3.1)
Italy	365 (5.4)	456 (3.0)	514 (3.8)	573 (4.2)	660 (6.5)
Japan	424 (4.2)	512 (3.7)	566 (3.0)	620 (2.7)	696 (4.5)
Latvia	403 (6.6)	473 (4.0)	519 (2.8)	562 (2.9)	623 (5.2)
Lithuania	392 (5.0)	469 (3.6)	522 (2.9)	571 (3.1)	642 (4.7)
Moldova, Rep. of	337 (11.3)	437 (5.0)	504 (5.5)	566 (6.0)	654 (9.3)
Morocco	217 (5.2)	299 (6.2)	360 (4.6)	420 (5.0)	500 (7.4)
Netherlands	440 (3.5)	494 (2.9)	531 (2.4)	566 (1.9)	617 (3.1)
New Zealand	349 (5.1)	437 (3.5)	494 (3.0)	551 (2.7)	629 (4.3)
Norway	315 (4.6)	398 (2.8)	451 (2.6)	501 (3.2)	568 (4.2)
Philippines	231 (6.6)	315 (4.5)	378 (3.9)	449 (8.9)	559 (14.3)
Russian Federation	381 (5.9)	455 (5.8)	511 (6.7)	5 617	

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

Exhibit D.2: Percentiles of Achievement in Applying Cognitive Domain

MATHEMATICS Grade 8

Countries					
Armenia	326 (4.5)	421 (3.8)	483 (4.1)	541 (3.7)	607 (2.2)
Australia	365 (7.6)	452 (5.4)	510 (5.4)	567 (6.2)	645 (6.8)
Bahrain	272 (2.0)	345 (2.9)	398 (2.6)	451 (2.5)	522 (3.7)
Belgium (Flemish)	402 (6.2)	494 (3.3)	543 (3.0)	587 (3.3)	642 (3.6)
Botswana	256	3 (2.2)			

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

Exhibit D.2: Percentiles of Achievement in Applying Cognitive Domain

MATHEMATICS Grade 4

Countries					
Armenia	326 (5.5)	407 (5.0)	461 (2.9)	516 (4.7)	598 (6.9)
Australia	350 (6.9)	434 (4.9)	491 (5.3)	545 (6.4)	626 (6.7)
Belgium (Flemish)	447 (4.6)	505 (1.7)			
			#•		Ð • ° ` ˆÀˆ• •‡XX b

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

Exhibit D.3: Percentiles of Achievement in Reasoning Cognitive Domain

MATHEMATICS Grade 8

Countries					95th Percentile
Armenia	293 (5.9)	402 (4.7)	472 (5.2)	541 (3.1)	622 (2.6)
Australia	375 (5.5)	461 (5.0)	520 (5.6)	572 (5.5)	642 (6.3)
Bahrain	297 (2.8)	374 (3.5)	426 (2.7)	477 (2.9)	545 (5.9)
Belgium (Flemish)	395 (5.5)	487 (3.7)	539 (2.3)	586 (3.3)	648 (3.3)
Botswana	204 (3.6)	292 (4.0)	353 (4.0)	415 (4.4)	502 (3.9)
Bulgaria	325 (6.8)	412 (4.4)	472 (5.2)	533 (4.7)	610 (5.4)
Chile	261 (5.5)	346 (5.0)	408 (4.0)	472 (4.3)	561 (4.2)
Chinese Taipei	414 (6.7)	514 (5.0)	581 (4.0)	642 (3.2)	721 (4.2)
Cyprus	308 (3.2)	399 (3.7)	458 (2.6)	516 (2.5)	593 (6.3)
Egypt	247 (6.8)	335 (3.4)	401 (3.4)	464 (5.2)	554 (5.6)
England	373 (5.9)	451 (6.6)	509 (6.5)	567 (6.5)	643 (7.4)
Estonia	399 (7.9)	472 (3.6)	524 (2.8)	574 (3.6)	643 (4.8)
Ghana	149 (4.7)	248 (4.0)	314 (5.5)	380 (5.9)	471 (5.4)
Hong Kong, SAR	436 (9.1)	522 (4.0)	574 (4.2)	621 (2.6)	684 (3.5)
Hungary	402 (6.0)	477 (3.6)	530 (3.0)	582 (4.7)	655 (6.9)
Indonesia	253 (7.5)	344 (5.5)	406 (6.1)	468 (4.7)	556 (7.6)
Iran, Islamic Rep. of	295 (5.0)	365 (3.2)	417 (2.9)	468 (2.7)	540 (2.4)
Israel	328 (4.1)	419 (4.1)	485 (5.4)	547 (3.8)	632 (5.4)
Italy	360 (8.1)	438 (4.1)	491 (3.2)	542 (3.3)	610 (5.4)
Japan	446 (7.1)	528 (3.1)	577 (1.5)	625 (3.3)	698 (5.1)
Jordan	295 (6.1)	377 (5.4)	435 (4.1)	490 (4.5)	566 (6.6)
Korea, Rep. of	441 (3.6)	530 (2.3)	585 (1.6)	638 (2.3)	712 (3.7)
Latvia	367 (6.4)	447 (4.2)	501 (3.4)	553 (3.7)	629 (5.7)
Lebanon	278 (6.9)	355 (3.1)	410 (3.4)	464 (3.7)	540 (4.4)
Lithuania	354 (5.1)	435 (4.3)	490 (2.6)	545 (2.7)	618 (3.2)
Macedonia, Rep. of	278 (8.3)	376 (4.5)	442 (4.9)	504 (5.6)	584 (4.6)
Malaysia	386 (2.9)	453 (2.8)	504 (4.9)	554 (4.1)	618 (3.9)
Moldova, Rep. of	309 (8.7)	395 (6.2)	456 (3.6)	513 (3.6)	590 (3.3)
Morocco	259 (7.4)	338 (4.2)	392 (3.2)	446 (4.5)	520 (8.2)
Netherlands	416 (6.2)	490 (4.8)	543 (4.0)	594 (5.3)	660 (8.3)
New Zealand	374 (5.2)	455 (5.5)	513 (6.8)	565 (8.1)	639 (7.0)
Norway	343 (6.0)	427 (3.6)	483 (3.0)	534 (2.7)	604 (3.3)
Palestinian Nat'l Auth.	259 (6.1)	346 (3.3)	405 (3.9)	465 (2.7)	545 (5.6)
Philippines	177 (6.1)	282 (5.4)	355 (7.4)	433 (6.2)	542 (8.0)
Romania	297 (8.7)	395 (4.6)	461 (5.1)	523 (5.5)	611 (4.8)
Russian Federation	365 (8.9)	443 (4.7)	497 (3.8)	551 (4.4)	623 (3.4)
Saudi Arabia	205 (7.9)	289 (6.1)	349 (3.9)	407 (3.7)	489 (6.5)
Scotland	375 (7.9)	459 (4.0)	518 (3.9)	570 (4.1)	638 (8.5)
Serbia	323 (6.0)	411 (3.4)	471 (3.5)	527 (2.7)	604 (5.5)
Singapore	424 (5.6)	528 (4.8)	591 (3.2)	645 (2.9)	717 (3.7)
Slovak Republic	360 (5.5)	450 (7.2)	506 (4.9)	560 (3.9)	636 (4.4)
Slovenia	372 (4.7)	445 (2.8)	495 (3.3)	543 (1.9)	610 (3.8)
South Africa	111 (8.0)	207 (3.6)	277 (3.4)	354 (7.1)	505 (13.7)
Sweden	364 (9.3)	453 (4.0)	510 (2.9)	567 (3.3)	642 (5.9)
Tunisia	282 (4.9)	351 (3.0)	400 (3.3)	447 (2.4)	516 (7.7)
United States	366 (5.0)	448 (4.1)	507 (4.8)	564 (2.9)	638 (4.5)
Benchmarking Participants					
Basque Country, Spain	371 (7.9)	446 (3.1)	496 (2.3)	545 (1.9)	610 (2.1)
Indiana State, US	376 (10.4)	451 (5.0)	502 (4.7)	555 (5.9)	627 (8.7)
Ontario Province, Can.	403 (9.6)	479 (3.6)	531 (3.8)	577 (3.5)	641 (4.8)
Quebec Province, Can.	429 (6.3)	494 (3.5)	540 (2.8)	583 (2.9)	646 (11.0)

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

Exhibit D.3: Percentiles of Achievement in Reasoning Cognitive Domain

MATHEMATICS Grade 4

Countries					
Armenia	305 (4.4)	387 (3.9)	445 (2.8)	503 (3.9)	588 (4.7)
Australia	373 (5.5)	455 (5.4)	510 (4.9)	561 (3.4)	633 (5.5)
Belgium (Flemish)	421 (4.3)	494 (3.3)	542 (3.0)	591 (3.5)	659 (4.3)
Chinese Taipei	422 (7.8)	513 (3.4)	568 (2.9)	618 (2.0)	691 (5.9)
Cyprus	359 (4.5)	460 (3.2)	519 (2.6)	574 (3.3)	653 (4.4)
England	390 (8.1)	479 (4.6)	539 (3.3)	598 (4.0)	678 (4.9)
Hong Kong, SAR	427 (6.4)	514 (4.4)	568 (4.4)	618 (3.0)	689 (3.9)
Hungary	379 (5.6)	467 (4.7)	528 (2.9)	584 (4.0)	661 (4.3)

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

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

Exhibit D.5: Standard Deviations of Achievement in Applying Cognitive Domain

MATHEMATICS Grade 4

Armenia	462 (3.2)	82 (1.9)	465 (3.2)	80 (2.0)	459 (3.7)	84 (2.2)
Australia	490 (3.8)	83 (2.0)	487 (4.3)	81 (2.2)	492 (4.4)	86 (2.5)
Belgium (Flemish)	546 (2.1)	60 (1.1)	544 (2.5)	59 (1.3)	548 (2.7)	62 (1.4)
Chinese Taipei	561 (1.9)	60 (465	4	59 (2.		

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

Exhibit D.6: Standard Deviations of Achievement in Reasoning Cognitive Domain

MATHEMATICS Grade 4

Armenia	445 (3.1)	86 (2.0)	449 (3.4)	84 (2.3)	442 (3.4)	87 (2.4)
Australia	507 (3.6)	79 (1.8)	507 (3.9)	77 (2.1)	507 (4.2)	80 (2.3)
Belgium (Flemish)	541 (2.2)	72 (1.5)	541 (2.6)	71 (2.0)	541 (2.8)	74 (1.7)
Chinese Taipei	563 (2.2)	81 (1.3)	565 (2.6)	76 (1.6)	562 (2.7)	85 (1.7)
Cyprus	516 (2.4)	86 (1.4)	515 (2.7)	83 (1.6)	517 (3.0)	88 (1.8)
England	537 (3.5)	87 (1.5)	539 (4.0)	85 (2.0)	536 (4.2)	90 (2.1)
Hong Kong, SAR	564 (3.7)	79 (1.4)	565 (4.0)	75 (1.7)	563 (3.8)	82 (2.0)
Hungary	524 (3.2)	86 (2.0)	525 (4.0)	85 (2.4)	524 (3.8)	87 (2.3)
Iran, Islamic Rep. of	400 (3.4)	80 (1.9)	406 (6.0)	79 (2.9)	396 (4.3)	81 (2.2)
IF . ǐ ľ d̬ʿ ẩ	EFg Kong, SAR	خ ס	ʿ d̬ Ɀ			
.)F(Eʹ ضƺ	HE . d̬ ľ d̬ʿ Ỗ EFB					