Asian Centre for Mathematics Education
East China Normal University, Shanghai, China

International Studies of Mathematics Education: Comparing the incomparables?

Frederick K. S. Leung
Kintoy Professor in Mathematics Education
The University of Hong Kong

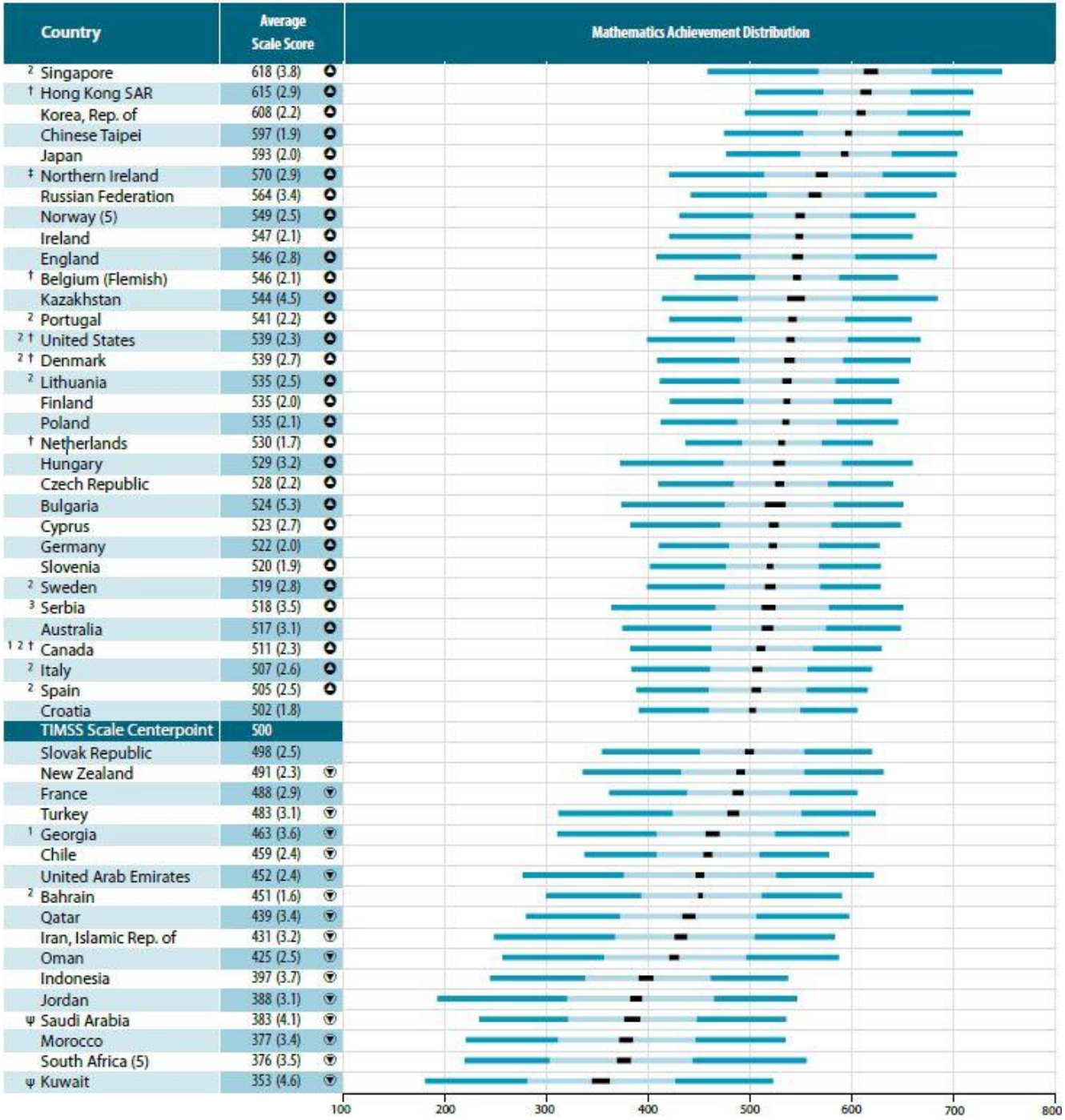
25 October 2018

Introduction

- In the past two decades, international studies of mathematics education such as TIMSS and PISA have attracted much attention in the mathematics education community and beyond.
- Some scholars however have queried the applicability of the results of these studies, pointing to issues of comparability due to the vast difference in context among the countries involved.
- In this presentation, both the strengths and limitations of international studies of mathematics education will be discussed.
- First let us look at the results of some of these studies

Grade 4

Mathematics



Country	Average Scale Score
[2] Singapore	618 (3.8)
† Hong Kong SAR	615 (2.9)
Korea, Rep. of	608 (2.2)
Chinese Taipei	597 (1.9)
Japan	593 (2.0)
‡ Northern Ireland	570 (2.9)
Russian Federation	564 (3.4)
Norway (5)	549 (2.5)
Ireland	547 (2.1)
England	546 (2.8)
† Belgium (Flemish)	546 (2.1)
Kazakhstan	544 (4.5)
[2] Portugal	541 (2.2)
[2] † United States	539 (2.3)
[2] † Denmark	539 (2.7)
[2] Lithuania	535 (2.5)
Finland	535 (2.0)
Poland	535 (2.1)
† Netherlands	530 (1.7)
Hungary	529 (3.2)
Czech Republic	528 (2.2)
Bulgaria	524 (5.3)
Cyprus	523 (2.7)
Germany	522 (2.0)
Slovenia	520 (1.9)
[2] Sweden	519 (2.8)
[3] Serbia	518 (3.5)
Australia	517 (3.1)
[1 2] † Canada	511 (2.3)
[2] Italy	507 (2.6)
[2] Spain	505 (2.5)
Croatia	502 (1.8)
TIMSS Scale Centerpoint	**500**
Slovak Republic	498 (2.5)
New Zealand	491 (2.3)
France	488 (2.9)
Turkey	483 (3.1)
[1] Georgia	463 (3.6)
Chile	459 (2.4)
United Arab Emirates	452 (2.4)
[2] Bahrain	451 (1.6)
Qatar	439 (3.4)
Iran, Islamic Rep. of	431 (3.2)
Oman	425 (2.5)
Indonesia	397 (3.7)
Jordan	388 (3.1)
ψ Saudi Arabia	383 (4.1)
Morocco	377 (3.4)
South Africa (5)	376 (3.5)
ψ Kuwait	353 (4.6)

Mathematics Achievement Distribution

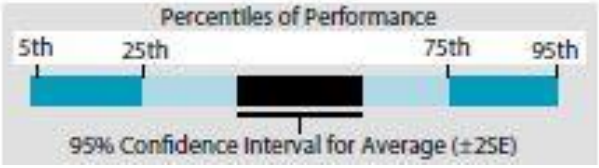
100 200 300 400 500 600 700 800

Grade 4 Mathematics (first 15 countries)

Country	Average Scale Score	Mathematics Achievement Distribution
[2] Singapore	618 (3.8) ▲	
† Hong Kong SAR	615 (2.9) ▲	
Korea, Rep. of	608 (2.2) ▲	
Chinese Taipei	597 (1.9) ▲	
Japan	593 (2.0) ▲	
‡ Northern Ireland	570 (2.9) ▲	
Russian Federation	564 (3.4) ▲	
Norway (5)	549 (2.5) ▲	
Ireland	547 (2.1) ▲	
England	546 (2.8) ▲	
† Belgium (Flemish)	546 (2.1) ▲	
Kazakhstan	544 (4.5) ▲	
[2] Portugal	541 (2.2) ▲	
[2] † United States	539 (2.3) ▲	
[2] † Denmark	539 (2.7) ▲	

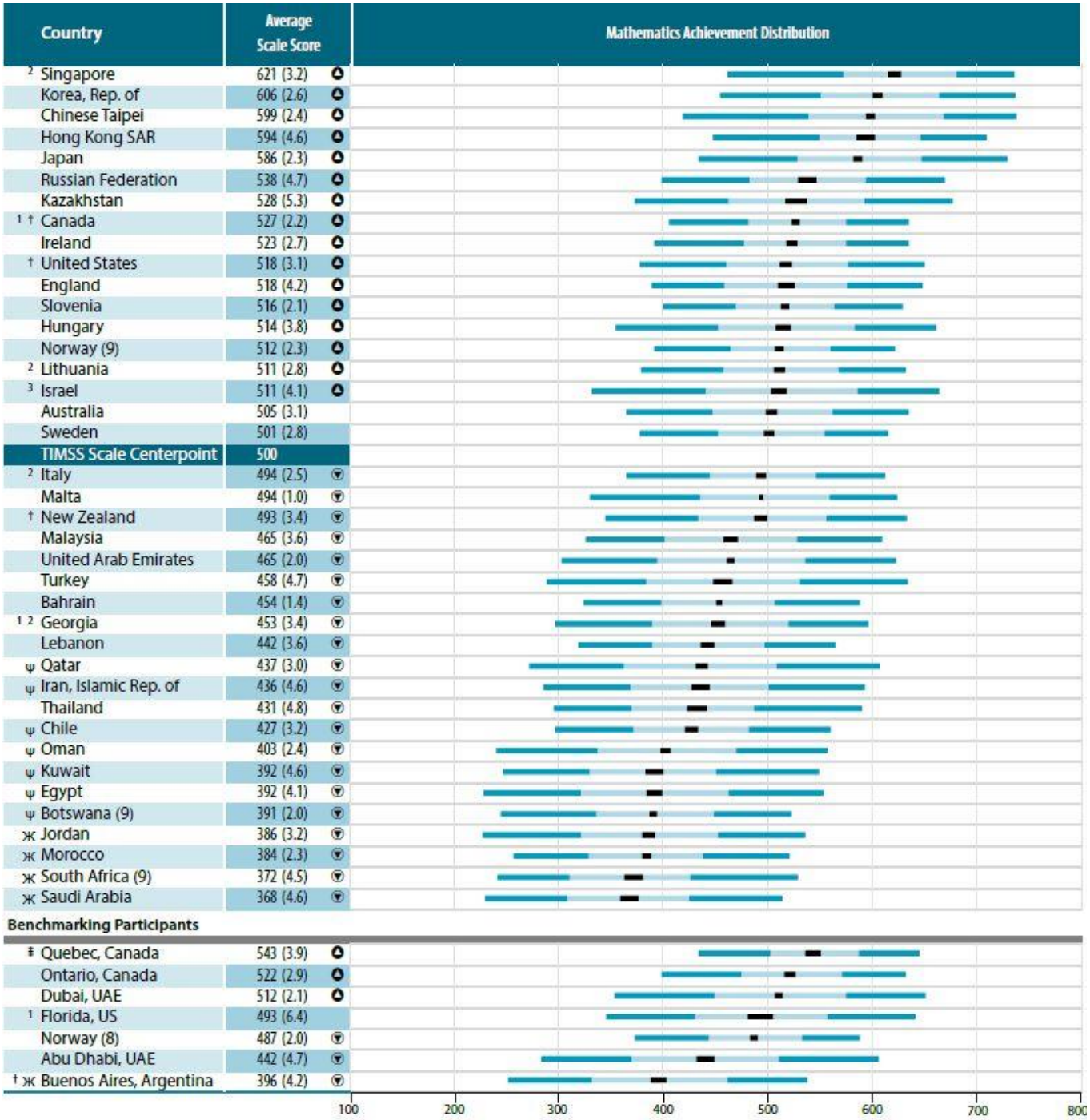
▲ Country average significantly higher than the centerpoint of the TIMSS 4th grade scale

▼ Country average significantly lower than the centerpoint of the TIMSS 4th grade scale



Grade 8

Mathematics



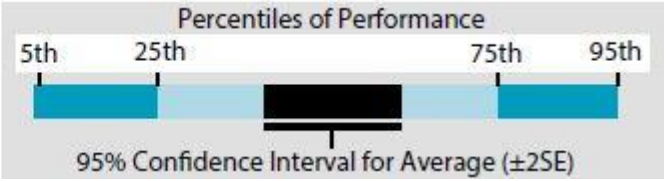
Country	Average Scale Score
[2] Singapore	621 (3.2) ▲
Korea, Rep. of	606 (2.6) ▲
Chinese Taipei	599 (2.4) ▲
Hong Kong SAR	594 (4.6) ▲
Japan	586 (2.3) ▲
Russian Federation	538 (4.7) ▲
Kazakhstan	528 (5.3) ▲
[1] † Canada	527 (2.2) ▲
Ireland	523 (2.7) ▲
† United States	518 (3.1) ▲
England	518 (4.2) ▲
Slovenia	516 (2.1) ▲
Hungary	514 (3.8) ▲
Norway (9)	512 (2.3) ▲
[2] Lithuania	511 (2.8) ▲
[3] Israel	511 (4.1) ▲
Australia	505 (3.1)
Sweden	501 (2.8)
TIMSS Scale Centerpoint	**500**
[2] Italy	494 (2.5) ▼
Malta	494 (1.0) ▼
† New Zealand	493 (3.4) ▼
Malaysia	465 (3.6) ▼
United Arab Emirates	465 (2.0) ▼
Turkey	458 (4.7) ▼
Bahrain	454 (1.4) ▼
[1] [2] Georgia	453 (3.4) ▼
Lebanon	442 (3.6) ▼
ψ Qatar	437 (3.0) ▼
ψ Iran, Islamic Rep. of	436 (4.6) ▼
Thailand	431 (4.8) ▼
ψ Chile	427 (3.2) ▼
ψ Oman	403 (2.4) ▼
ψ Kuwait	392 (4.6) ▼
ψ Egypt	392 (4.1) ▼
ψ Botswana (9)	391 (2.0) ▼
⁎ Jordan	386 (3.2) ▼
⁎ Morocco	384 (2.3) ▼
⁎ South Africa (9)	372 (4.5) ▼
⁎ Saudi Arabia	368 (4.6) ▼
Benchmarking Participants	
ǂ Quebec, Canada	543 (3.9) ▲
Ontario, Canada	522 (2.9) ▲
Dubai, UAE	512 (2.1) ▲
[1] Florida, US	493 (6.4)
Norway (8)	487 (2.0) ▼
Abu Dhabi, UAE	442 (4.7) ▼
† ⁎ Buenos Aires, Argentina	396 (4.2) ▼

Grade 8 Mathematics (first 15 countries)

Country	Average Scale Score	Mathematics Achievement Distribution
[2] Singapore	621 (3.2) ▲	
Korea, Rep. of	606 (2.6) ▲	
Chinese Taipei	599 (2.4) ▲	
Hong Kong SAR	594 (4.6) ▲	
Japan	586 (2.3) ▲	
Russian Federation	538 (4.7) ▲	
Kazakhstan	528 (5.3) ▲	
[1] † Canada	527 (2.2) ▲	
Ireland	523 (2.7) ▲	
† United States	518 (3.1) ▲	
England	518 (4.2) ▲	
Slovenia	516 (2.1) ▲	
Hungary	514 (3.8) ▲	
Norway (9)	512 (2.3) ▲	
[2] Lithuania	511 (2.8) ▲	

▲ Country average significantly higher than the centerpoint of the TIMSS 8th grade scale

▼ Country average significantly lower than the centerpoint of the TIMSS 8th grade scale



PISA 2015

Figure I.5.1 ▪ **Comparing countries' and economies' performance in mathematics**

Statistically significantly **above** the OECD average
Not statistically significantly different from the OECD average
Statistically significantly **below** the OECD average

Mean score	Comparison country/ economy	Countries and economies whose mean score is NOT statistically significantly different from the comparison country's/economy's score
564	Singapore	
548	Hong Kong (China)	Macao (China), Chinese Taipei
544	Macao (China)	Hong Kong (China), Chinese Taipei
542	Chinese Taipei	Hong Kong (China), Macao (China), B-S-J-G (China)
532	Japan	B-S-J-G (China), Korea
531	B-S-J-G (China)	Chinese Taipei, Japan, Korea, Switzerland
524	Korea	Japan, B-S-J-G (China), Switzerland, Estonia, Canada
521	Switzerland	B-S-J-G (China), Korea, Estonia, Canada
520	Estonia	Korea, Switzerland, Canada
516	Canada	Korea, Switzerland, Estonia, Netherlands, Denmark, Finland
512	Netherlands	Canada, Denmark, Finland, Slovenia, Belgium, Germany
511	Denmark	Canada, Netherlands, Finland, Slovenia, Belgium, Germany
511	Finland	Canada, Netherlands, Denmark, Slovenia, Belgium, Germany
510	Slovenia	Netherlands, Denmark, Finland, Belgium, Germany
507	Belgium	Netherlands, Denmark, Finland, Slovenia, Germany, Poland, Ireland, Norway
506	Germany	Netherlands, Denmark, Finland, Slovenia, Belgium, Poland, Ireland, Norway
504	Poland	Belgium, Germany, Ireland, Norway
504	Ireland	Belgium, Germany, Poland, Norway, Viet Nam
502	Norway	Belgium, Germany, Poland, Ireland, Austria, Viet Nam
497	Austria	Norway, New Zealand, Viet Nam, Russia, Sweden, Australia, France, United Kingdom, Czech Republic, Portugal, Italy
495	New Zealand	Austria, Viet Nam, Russia, Sweden, Australia, France, United Kingdom, Czech Republic, Portugal, Italy
495	Viet Nam	Ireland, Norway, Austria, New Zealand, Russia, Sweden, Australia, France, United Kingdom, Czech Republic, Portugal, Italy, Iceland, Spain, Luxembourg
494	Russia	Austria, New Zealand, Viet Nam, Sweden, Australia, France, United Kingdom, Czech Republic, Portugal, Italy, Iceland
494	Sweden	Austria, New Zealand, Viet Nam, Russia, Australia, France, United Kingdom, Czech Republic, Portugal, Italy, Iceland
494	Australia	Austria, New Zealand, Viet Nam, Russia, Sweden, France, United Kingdom, Czech Republic, Portugal, Italy
493	France	Austria, New Zealand, Viet Nam, Russia, Sweden, Australia, United Kingdom, Czech Republic, Portugal, Italy, Iceland
492	United Kingdom	Austria, New Zealand, Viet Nam, Russia, Sweden, Australia, France, Czech Republic, Portugal, Italy, Iceland
492	Czech Republic	Austria, New Zealand, Viet Nam, Russia, Sweden, Australia, France, United Kingdom, Portugal, Italy, Iceland
492	Portugal	Austria, New Zealand, Viet Nam, Russia, Sweden, Australia, France, United Kingdom, Czech Republic, Italy, Iceland, Spain
490	Italy	Austria, New Zealand, Viet Nam, Russia, Sweden, Australia, France, United Kingdom, Czech Republic, Portugal, Iceland, Spain, Luxembourg
488	Iceland	Viet Nam, Russia, Sweden, France, United Kingdom, Czech Republic, Portugal, Italy, Spain, Luxembourg
486	Spain	Viet Nam, Portugal, Italy, Iceland, Luxembourg, Latvia
486	Luxembourg	Viet Nam, Italy, Iceland, Spain, Latvia
482	Latvia	Spain, Luxembourg, Malta, Lithuania, Hungary
479	Malta	Latvia, Lithuania, Hungary, Slovak Republic
478	Lithuania	Latvia, Malta, Hungary, Slovak Republic
477	Hungary	Latvia, Malta, Lithuania, Slovak Republic, Israel, United States
475	Slovak Republic	Malta, Lithuania, Hungary, Israel, United States
470	Israel	Hungary, Slovak Republic, United States, Croatia, CABA (Argentina)
470	United States	Hungary, Slovak Republic, Israel, Croatia, CABA (Argentina)
464	Croatia	Israel, United States, CABA (Argentina)
456	CABA (Argentina)	Israel, United States, Croatia, Greece, Romania, Bulgaria
454	Greece	CABA (Argentina), Romania
444	Romania	CABA (Argentina), Greece, Bulgaria, Cyprus[1]
441	Bulgaria	CABA (Argentina), Romania, Cyprus[1]
437	Cyprus[1]	Romania, Bulgaria
427	United Arab Emirates	Chile, Turkey
423	Chile	United Arab Emirates, Turkey, Moldova, Uruguay, Montenegro, Trinidad and Tobago, Thailand
420	Turkey	United Arab Emirates, Chile, Moldova, Uruguay, Montenegro, Trinidad and Tobago, Thailand, Albania
420	Moldova	Chile, Turkey, Uruguay, Montenegro, Trinidad and Tobago, Thailand, Albania
418	Uruguay	Chile, Turkey, Moldova, Montenegro, Trinidad and Tobago, Thailand, Albania
418	Montenegro	Chile, Turkey, Moldova, Uruguay, Trinidad and Tobago, Thailand, Albania
417	Trinidad and Tobago	Chile, Turkey, Moldova, Uruguay, Montenegro, Thailand, Albania
415	Thailand	Chile, Turkey, Moldova, Uruguay, Montenegro, Trinidad and Tobago, Albania
413	Albania	Turkey, Moldova, Uruguay, Montenegro, Trinidad and Tobago, Thailand, Mexico
408	Mexico	Albania, Georgia
404	Georgia	Mexico, Qatar, Costa Rica, Lebanon
402	Qatar	Georgia, Costa Rica, Lebanon
400	Costa Rica	Georgia, Qatar, Lebanon
396	Lebanon	Georgia, Qatar, Costa Rica, Colombia
390	Colombia	Lebanon, Peru, Indonesia
387	Peru	Colombia, Indonesia, Jordan
386	Indonesia	Colombia, Peru, Jordan
380	Jordan	Peru, Indonesia, Brazil
377	Brazil	Jordan, FYROM
371	FYROM	Brazil, Tunisia
367	Tunisia	FYROM, Kosovo, Algeria
362	Kosovo	Tunisia, Algeria
360	Algeria	Tunisia, Kosovo
328	Dominican Republic	

Figure I.5.1 ■ **Comparing countries' and economies' performance in mathematics**

Statistically significantly **above** the OECD average
Not statistically significantly different from the OECD average
Statistically significantly **below** the OECD average

Mean score	Comparison country/ economy	Countries and economies whose mean score is NOT statistically significantly different from the comparison country's/economy's score
564	Singapore	
548	Hong Kong (China)	Macao (China), Chinese Taipei
544	Macao (China)	Hong Kong (China), Chinese Taipei
542	Chinese Taipei	Hong Kong (China), Macao (China), B-S-J-G (China)
532	Japan	B-S-J-G (China), Korea
531	B-S-J-G (China)	Chinese Taipei, Japan, Korea, Switzerland
524	Korea	Japan, B-S-J-G (China), Switzerland, Estonia, Canada
521	Switzerland	B-S-J-G (China), Korea, Estonia, Canada
520	Estonia	Korea, Switzerland, Canada
516	Canada	Korea, Switzerland, Estonia, Netherlands, Denmark, Finland
512	Netherlands	Canada, Denmark, Finland, Slovenia, Belgium, Germany
511	Denmark	Canada, Netherlands, Finland, Slovenia, Belgium, Germany
511	Finland	Canada, Netherlands, Denmark, Slovenia, Belgium, Germany
510	Slovenia	Netherlands, Denmark, Finland, Belgium, Germany
507	Belgium	Netherlands, Denmark, Finland, Slovenia, Germany, Poland, Ireland, Norway
506	Germany	Netherlands, Denmark, Finland, Slovenia, Belgium, Poland, Ireland, Norway
504	Poland	Belgium, Germany, Ireland, Norway
504	Ireland	Belgium, Germany, Poland, Norway, Viet Nam
502	Norway	Belgium, Germany, Poland, Ireland, Austria, Viet Nam
497	Austria	Norway, New Zealand, Viet Nam, Russia, Sweden, Australia, France, United Kingdom, Czech Republic, Portugal, Italy
495	New Zealand	Austria, Viet Nam, Russia, Sweden, Australia, France, United Kingdom, Czech Republic, Portugal, Italy
495	Viet Nam	Ireland, Norway, Austria, New Zealand, Russia, Sweden, Australia, France, United Kingdom, Czech Republic, Portugal, Italy, Iceland, Spain, Luxembourg
494	Russia	Austria, New Zealand, Viet Nam, Sweden, Australia, France, United Kingdom, Czech Republic, Portugal, Italy, Iceland
494	Sweden	Austria, New Zealand, Viet Nam, Russia, Australia, France, United Kingdom, Czech Republic, Portugal, Italy, Iceland
494	Australia	Austria, New Zealand, Viet Nam, Russia, Sweden, France, United Kingdom, Czech Republic, Portugal, Italy
493	France	Austria, New Zealand, Viet Nam, Russia, Sweden, Australia, United Kingdom, Czech Republic, Portugal, Italy, Iceland
492	United Kingdom	Austria, New Zealand, Viet Nam, Russia, Sweden, Australia, France, Czech Republic, Portugal, Italy, Iceland
492	Czech Republic	Austria, New Zealand, Viet Nam, Russia, Sweden, Australia, France, United Kingdom, Portugal, Italy, Iceland
492	Portugal	Austria, New Zealand, Viet Nam, Russia, Sweden, Australia, France, United Kingdom, Czech Republic, Italy, Iceland, Spain
490	Italy	Austria, New Zealand, Viet Nam, Russia, Sweden, Australia, France, United Kingdom, Czech Republic, Portugal, Iceland, Spain, Luxembourg
488	Iceland	Viet Nam, Russia, Sweden, France, United Kingdom, Czech Republic, Portugal, Italy, Spain, Luxembourg

Mean score	Comparison country/ economy	Countries and economies whose mean score is NOT statistically significantly
564	Singapore	
548	Hong Kong (China)	Macao (China), Chinese Taipei
544	Macao (China)	Hong Kong (China), Chinese Taipei
542	Chinese Taipei	Hong Kong (China), Macao (China), B-S-J-G (China)
532	Japan	B-S-J-G (China), Korea
531	B-S-J-G (China)	Chinese Taipei, Japan, Korea, Switzerland
524	Korea	Japan, B-S-J-G (China), Switzerland, Estonia, Canada
521	Switzerland	B-S-J-G (China), Korea, Estonia, Canada
520	Estonia	Korea, Switzerland, Canada
516	Canada	Korea, Switzerland, Estonia, Netherlands, Denmark, Finland
512	Netherlands	Canada, Denmark, Finland, Slovenia, Belgium, Germany
511	Denmark	Canada, Netherlands, Finland, Slovenia, Belgium, Germany
511	Finland	Canada, Netherlands, Denmark, Slovenia, Belgium, Germany
510	Slovenia	Netherlands, Denmark, Finland, Belgium, Germany
507	Belgium	Netherlands, Denmark, Finland, Slovenia, Germany, Poland, Ireland, Norway
506	Germany	Netherlands, Denmark, Finland, Slovenia, Belgium, Poland, Ireland, Norway
504	Poland	Belgium, Germany, Ireland, Norway
504	Ireland	Belgium, Germany, Poland, Norway, Viet Nam
502	Norway	Belgium, Germany, Poland, Ireland, Austria, Viet Nam

	Mean score	95% confidence interval
Singapore	564	561 - 567
Hong Kong (China)	548	542 - 554
Quebec (Canada)[1]	544	535 - 553
Macao (China)	544	542 - 546
Chinese Taipei	542	536 - 548
Japan	532	527 - 538
B-S-J-G (China)	531	522 - 541
Korea	524	517 - 531
British Columbia (Canada)	522	512 - 531
Flemish community (Belgium)	521	517 - 526
Switzerland	521	516 - 527
Estonia	520	516 - 524
Bolzano (Italy)	518	505 - 531
Navarre (Spain)	518	503 - 533
Trento (Italy)	516	511 - 521
Canada	516	511 - 520
Netherlands	512	508 - 517

What do these international studies purport to do?

Goals of TIMSS 2015

"TIMSS has the goal of helping countries make informed decisions about how to improve teaching and learning in mathematics and science. With its strong curricular focus and emphasis on policy relevant information about the home, school, and classroom contexts for learning, TIMSS is a valuable tool that countries can use to evaluate achievement goals and standards and monitor students' achievement trends in an international context."

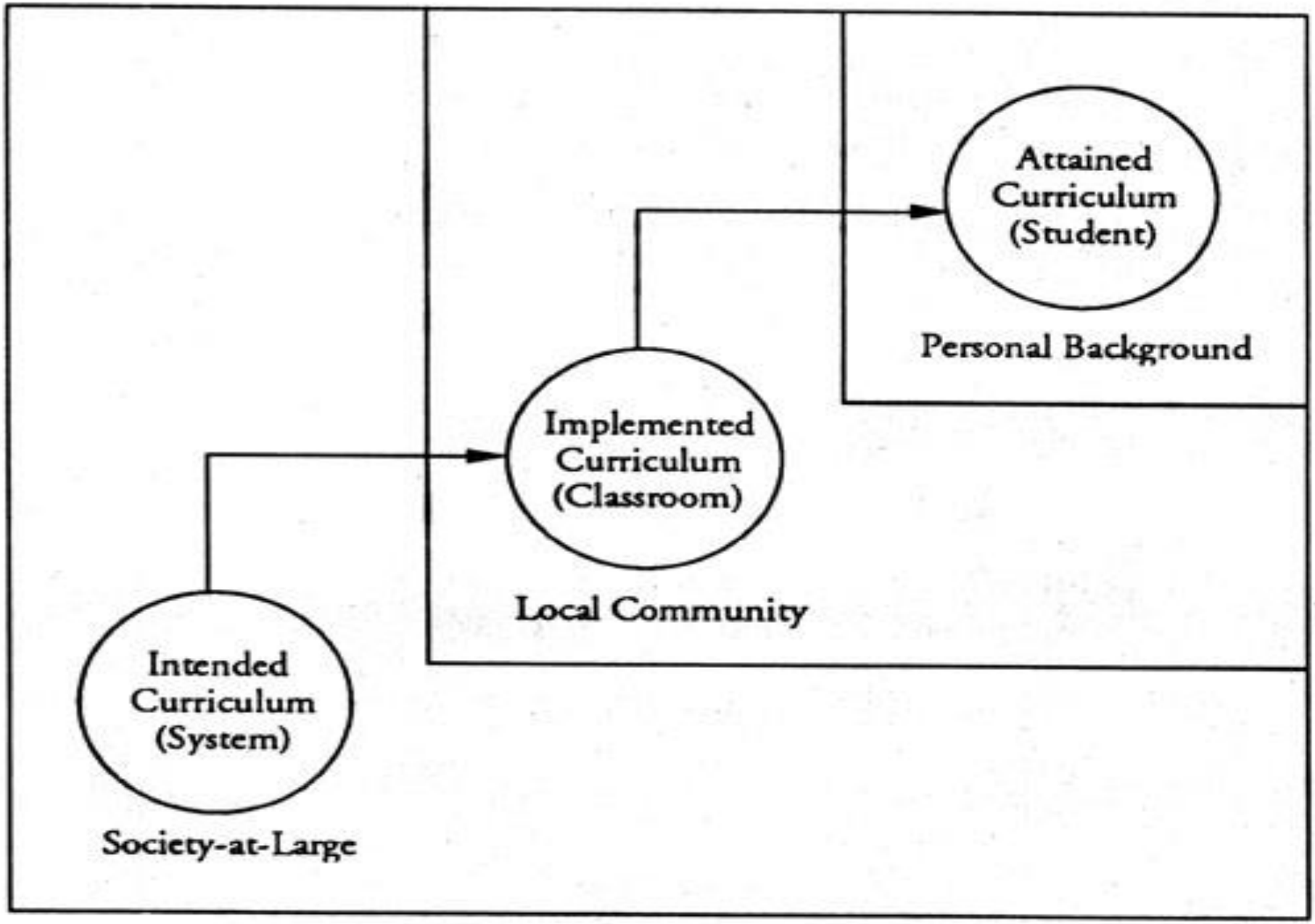


Figure 1. Conceptual Framework for TIMSS.

Goals of PISA

- PISA claims to assess “What is important for citizens to know and be able to do?” (literacy)
- “PISA assesses the extent to which 15-year-old students, near the end of their compulsory education, have acquired key knowledge and skills that are essential for full participation in modern societies.
- The findings allow policy makers around the world to gauge the knowledge and skills of students in their own countries in comparison with those in other countries, set policy targets against measurable goals achieved by other education systems, and learn from policies and practices applied elsewhere.
- PISA is different from other international assessments in its **policy orientation**, …”

The importance of International Studies

- Many variables within a country are uniform and cannot be manipulated, and to study the impact of those variables on student achievement, we have to collect data in different countries, where the variables differ
- It may be impractical and unethical to manipulate some variables within a country
- International studies of mathematics education – using the world as “a natural educational research laboratory”
- In a laboratory, the conditions are made uniform

1. **Uniform curriculum framework based on consensus of participating countries**

Content Aspect

Detailed Categories

1.1 Numbers

1.1.1 Whole Numbers

1.1.1.1 Meaning
[This includes the uses of numbers; place value and numeration; ordering and comparing numbers]

1.1.1.2 Operations
[Including addition; subtraction; multiplication; division; mixed operations]

1.1.1.3 Properties of operations
[Including commutative property, distributive property, etc.]

1.1.2 Fractions and Decimals

1.1.2.1 Common fractions
[Includes meaning and representation of common fractions; computation with common fractions and mixed numbers]

1.1.2.2 Decimal fractions
[Includes meaning and representation of decimals; computation with decimals]

1.1.2.3 Relationships between common and decimal fractions
[Including conversion to equivalent forms; ordering of fractions and decimals]

1.1.2.4 Percentage

Performance Expectations Aspect

Detailed Categories

2.1 Knowing

2.1.1 Representing
[This includes demonstrating knowledge of a nonverbal mathematical representation of a mathematical object or procedure either by selection or by construction, either formal or informal. Representations might be concrete, pictorial, graphical, algebraic, etc.

2.1.2 Recognizing equivalents
[This includes selecting or constructing mathematically equivalent objects (e.g., equivalent common and decimal fractions; equivalent trigonometric functions and power series; equivalent representations of concepts [e.g., place value]; equivalent axiomatic systems; etc.).

2.1.3 Recalling mathematical objects and properties
[fitting given conditions]

2.2 Using Routine Procedures

2.2.1 Using equipment
[using instruments; using calculators and computers]

2.2.2 Performing routine procedures
[This includes counting and routine computations; graphing; transforming one mathematical object into another by some formal process; e.g., multiplying by a

Perspectives Aspect

Detailed Categories

3.1 Attitudes toward Science, Mathematics, and Technology

Encouraging positive attitudes toward science, mathematics, and technology.

3.2 Careers involving Science, Mathematics, and Technology

3.2.1 Promoting careers in science, mathematics, and technology.

3.2.2 Promoting the importance of science, mathematics, and technology in non-technical careers

3.3 Participation in Science and Mathematics by Underrepresented Groups

Encouraging all types of students to study and use science, mathematics, and technology. Examples of groups that could be targeted are women or racial and ethnic minorities.

3.4 Science, Mathematics, and Technology to Increase Interest

Promoting interest and increasing understanding of topics in science, mathematics, and technology by using experiences that are common to students or popular or intriguing

Content and Cognitive Domains in TIMSS 2015

Content Domains for Grade 4
Number (50%)
Geometric Shapes and Measures (35%)
Data Display (15%)

Content Domains for Grade 8
Number (30%)
Algebra (30%)
Geometry (20%)
Data and Chance (20%)

Cognitive Domains for Grades 4 and 8
Knowing (35-40%)
Applying (40%)
Reasoning (20-25%)

2. Items based on the curricula of participating countries: TIMSS 2015 Test specification grid

Number of Mathematics Items of Each Type and Score Points, by Reporting Category, Grade 8
Score points are shown in parentheses. Source: Mullis et al. (2016), p. 327.

Reporting Category		Multiple-Choice	Constructed-Response	Total Items
Content domain	Number	29 (29)	35 (41)	64 (70)
	Algebra	35 (35)	27 (30)	62 (65)
	Geometry	22 (22)	21 (25)	43 (47)
	Data and chance	29 (31)	14 (16)	43 (47)
	Total	115 (117)	97 (112)	212 (229)
Cognitive domain	Knowing	50 (50)	19 (20)	69 (70)
	Applying	48 (48)	47 (55)	95 (103)
	Reasoning	17 (19)	31 (37)	48 (56)
	Total	115 (117)	97 (112)	212 (229)

3. Students carefully sampled (PPS sampling)

Total MOS: 72154 Sampling Interval: 481

School Sample: 150 Random Start: 236

School Code	School MOS	Cumulative MOS	Sample
917740	232	232	
875870	217	449	✓
924942	187	636	R
893204	161	797	✓
952774	159	956	R
806290	237	1193	
161758	206	1399	✓
357056	85	1484	R
997650	150	1634	
778732	141	1775	✓
216873	128	1903	R
336426	211	2114	
149238	232	2346	✓
		Total: 72154	

4. Rigorous data analysis

- Scaling of items: IRT scaling
- Scaling across countries, across cycles
- Use of plausible values (PV)
- IDB analyzer

5. Quality assurance measures

- Standardized data collection method
- Detailed instructions for data collection
- Training programs for data collection and scoring
- Double scoring and data re-entry for 1/3 of the data
- Translation of test items
  - 2 independent translators, back translation
  - Item statistics
- Data cleaning, item statistics
- National and international quality control monitors

Have we established comparability?

- Given the rigorous methodology adopted in these studies, could we legitimately compare students' achievements across countries?
- The major purpose of these international studies is to identify variables which explain achievement, for **informing policy decisions**
- But, can the "instructional, curricular, and resource related variables" really explain student achievement, and hence the results able to guide "educational decision making and practice in the areas of mathematics and science"? (TIMSS 2007 goals)

Policy implications of these studies?

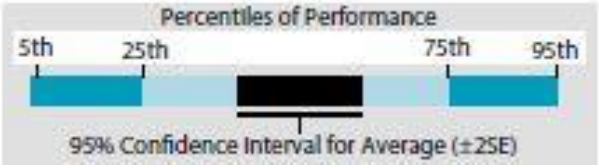
- Which countries do well, and why?
- For example, a pattern emerged: East Asian countries (Chinese Taipei, Hong Kong, Japan, Korea, and Singapore) performed very well in mathematics in these studies
- Is there anything that we can learn from this phenomenon?
- Following the purposes of these studies mentioned above, let us look at what these high achieving countries have In common to see whether we can identify shared "instructional, curricular, and resource related variables" that can be used to explain student achievement, and hence inform policy and practice

Grade 4 Mathematics (first 15 countries)

Country	Average Scale Score	Mathematics Achievement Distribution
[2] Singapore	618 (3.8) ▲	
† Hong Kong SAR	615 (2.9) ▲	
Korea, Rep. of	608 (2.2) ▲	
Chinese Taipei	597 (1.9) ▲	
Japan	593 (2.0) ▲	
‡ Northern Ireland	570 (2.9) ▲	
Russian Federation	564 (3.4) ▲	
Norway (5)	549 (2.5) ▲	
Ireland	547 (2.1) ▲	
England	546 (2.8) ▲	
† Belgium (Flemish)	546 (2.1) ▲	
Kazakhstan	544 (4.5) ▲	
[2] Portugal	541 (2.2) ▲	
[2] † United States	539 (2.3) ▲	
[2] † Denmark	539 (2.7) ▲	

▲ Country average significantly higher than the centerpoint of the TIMSS 4th grade scale

▼ Country average significantly lower than the centerpoint of the TIMSS 4th grade scale

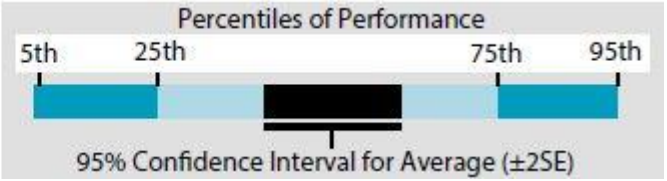


Grade 8 Mathematics (first 15 countries)

Country	Average Scale Score	Mathematics Achievement Distribution
Singapore	621 (3.2)	
Korea, Rep. of	606 (2.6)	
Chinese Taipei	599 (2.4)	
Hong Kong SAR	594 (4.6)	
Japan	586 (2.3)	
Russian Federation	538 (4.7)	
Kazakhstan	528 (5.3)	
1 † Canada	527 (2.2)	
Ireland	523 (2.7)	
† United States	518 (3.1)	
England	518 (4.2)	
Slovenia	516 (2.1)	
Hungary	514 (3.8)	
Norway (9)	512 (2.3)	
2 Lithuania	511 (2.8)	

▲ Country average significantly higher than the centerpoint of the TIMSS 8th grade scale

▼ Country average significantly lower than the centerpoint of the TIMSS 8th grade scale

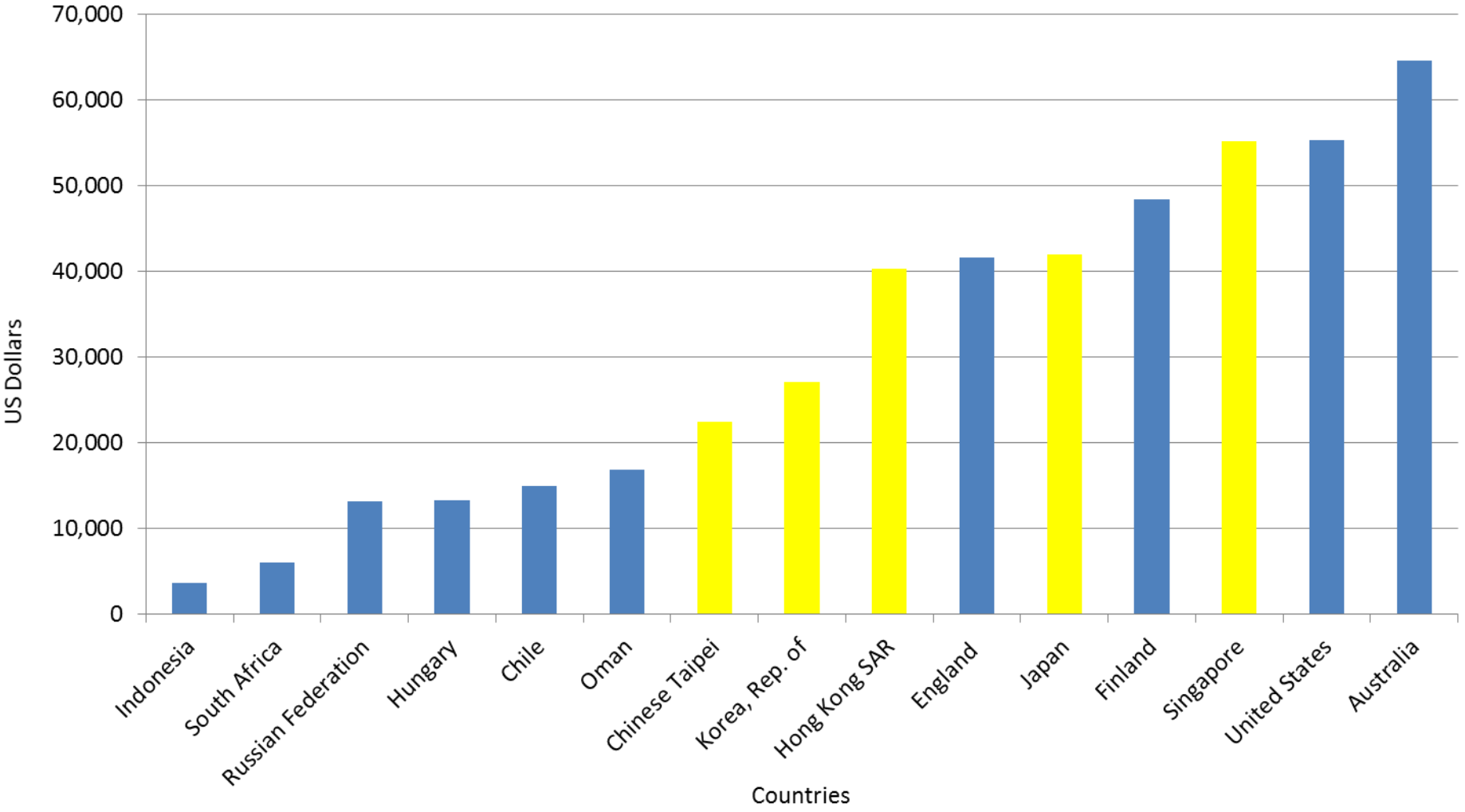


PISA 2015

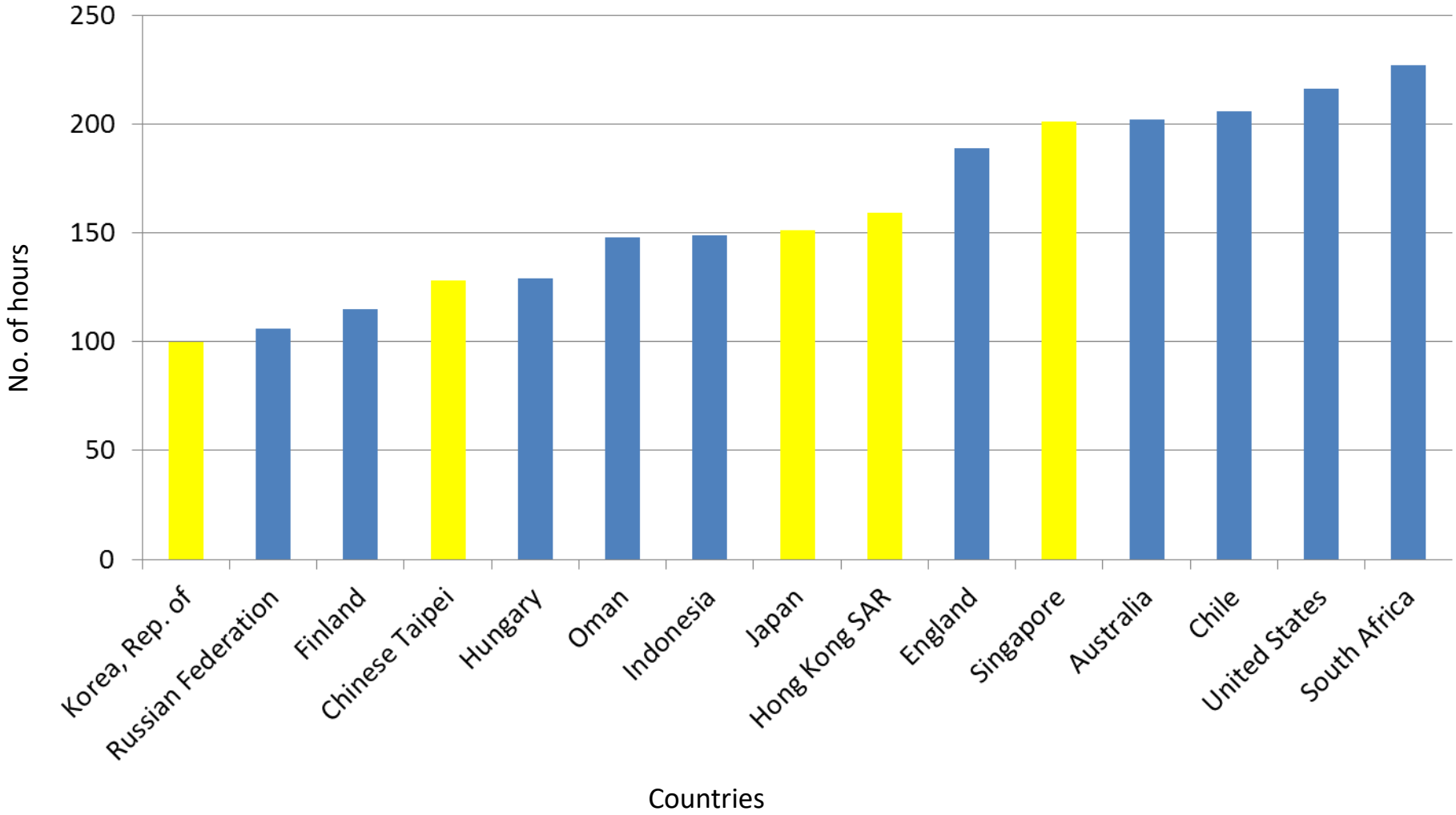
	Mean score	95% confidence interval
Singapore	564	561 - 567
Hong Kong (China)	548	542 - 554
Quebec (Canada)[1]	544	535 - 553
Macao (China)	544	542 - 546
Chinese Taipei	542	536 - 548
Japan	532	527 - 538
B-S-J-G (China)	531	522 - 541
Korea	524	517 - 531
British Columbia (Canada)	522	512 - 531
Flemish community (Belgium)	521	517 - 526
Switzerland	521	516 - 527
Estonia	520	516 - 524
Bolzano (Italy)	518	505 - 531
Navarre (Spain)	518	503 - 533
Trento (Italy)	516	511 - 521
Canada	516	511 - 520
Netherlands	512	508 - 517

Background variables

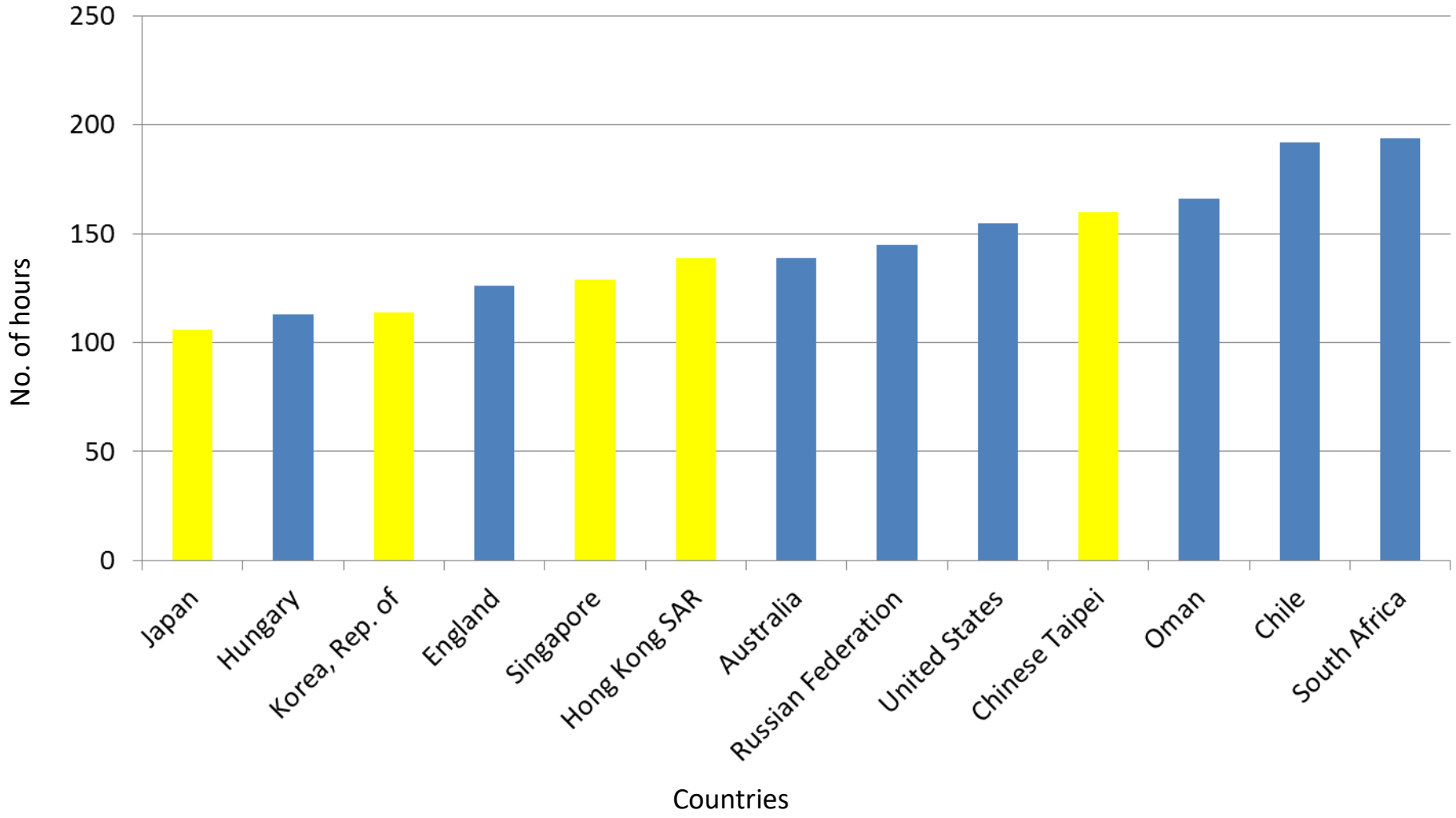
GNI per capita (US$)



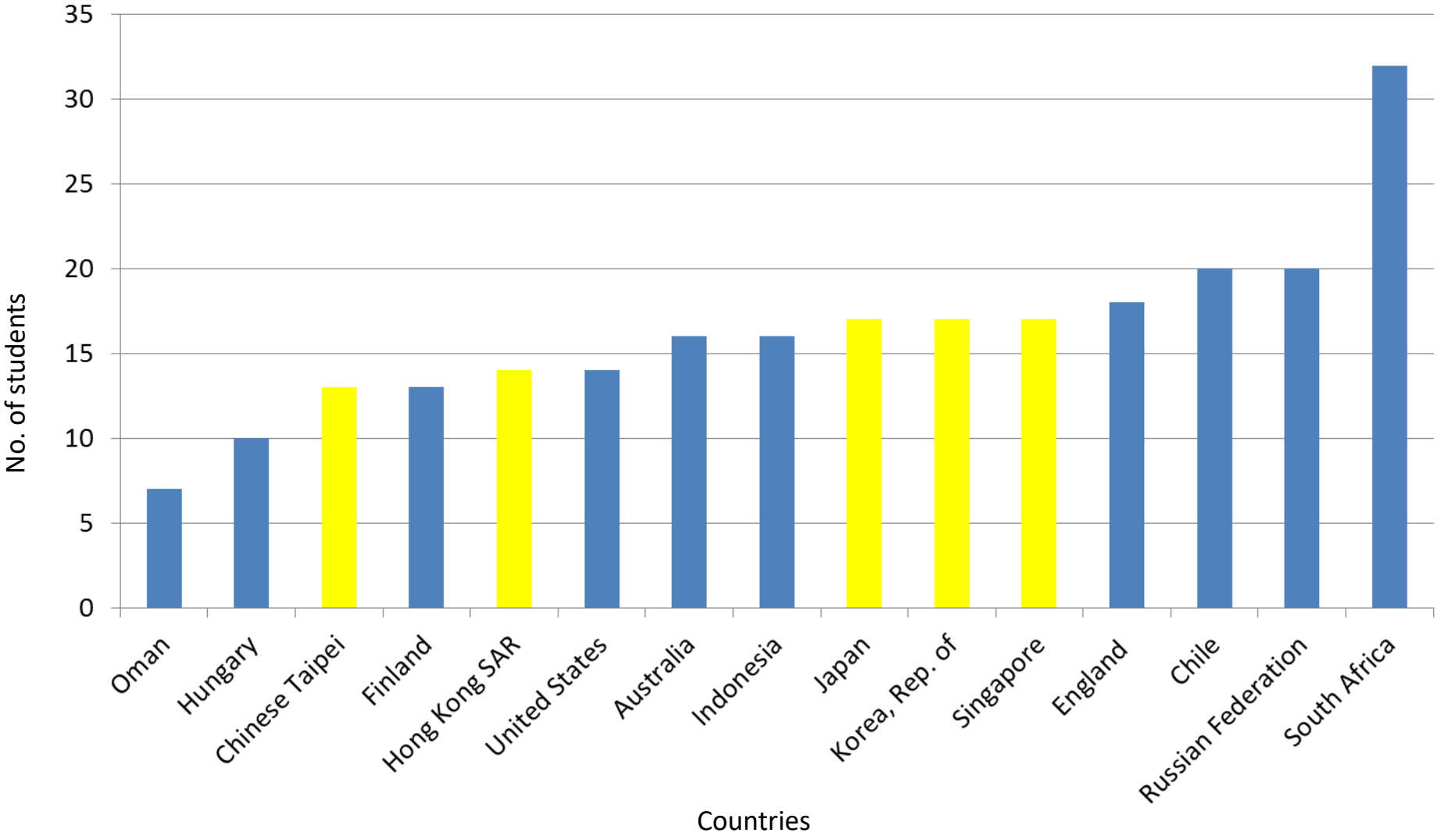
Mathematics Instructional Time (G.4 Int'l Avg. = 157 hours)



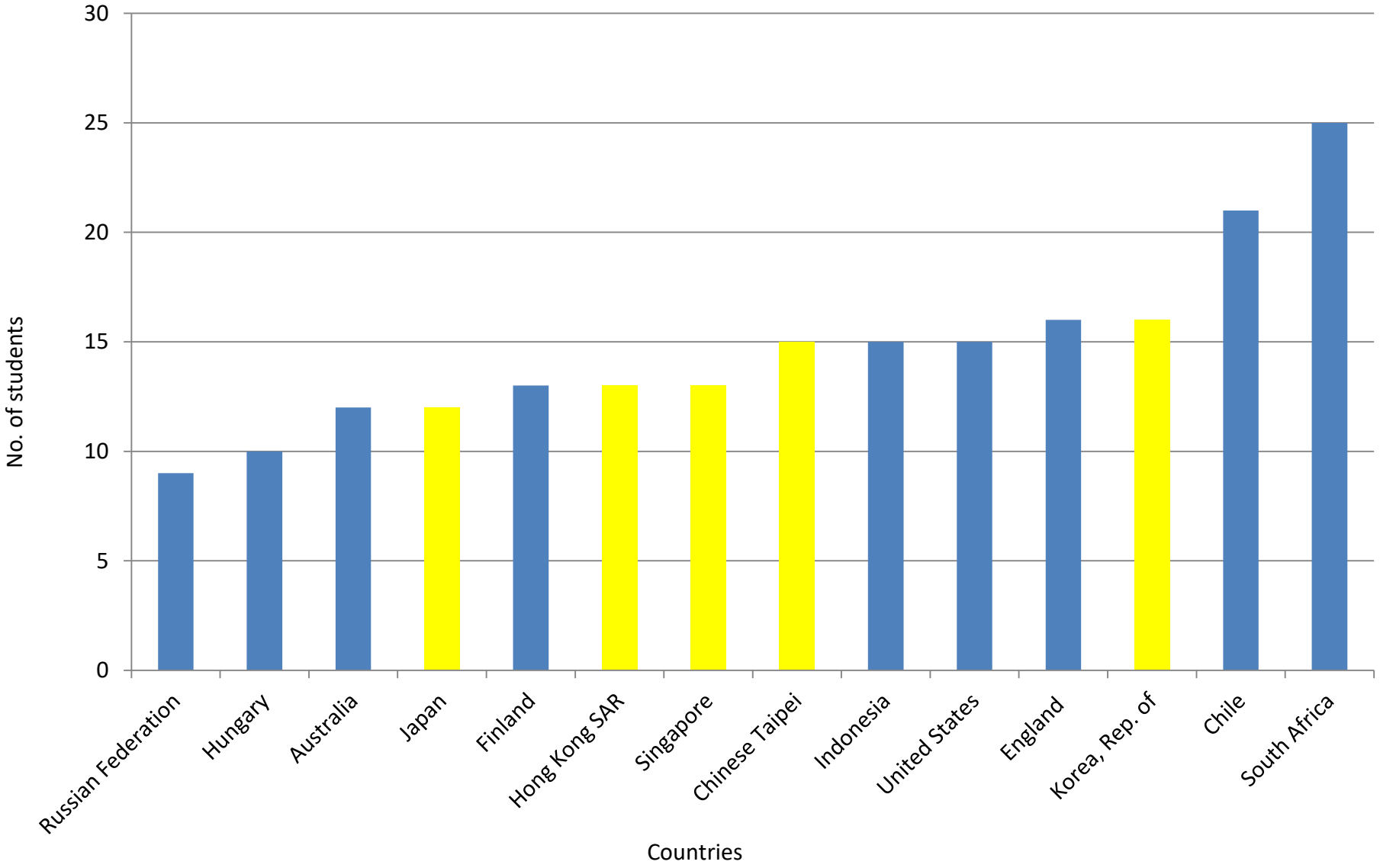
Mathematics Instructional Time (G.8 Int'l Avg. = 138 hours)



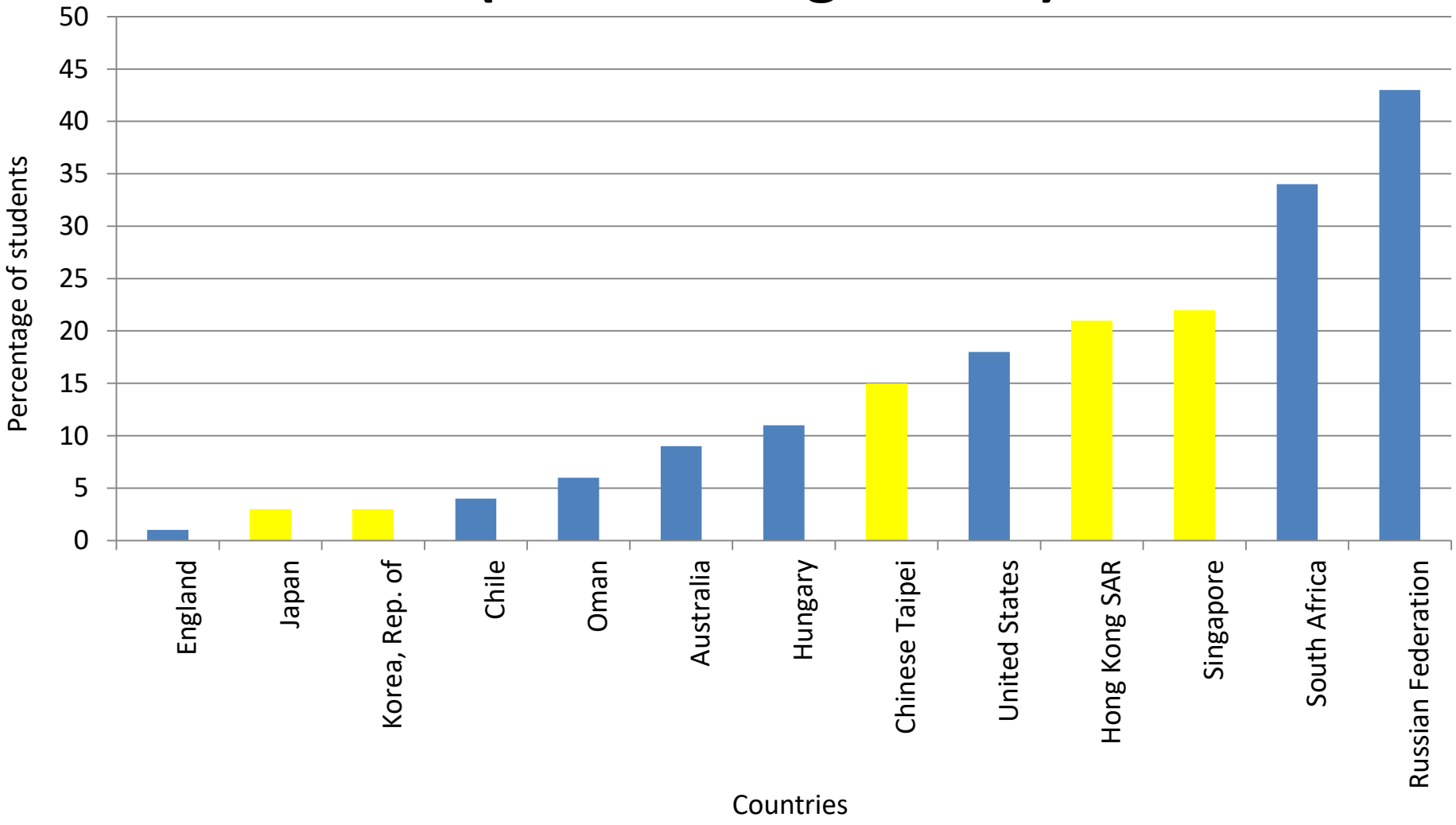
Primary Pupil-Teacher Ratio



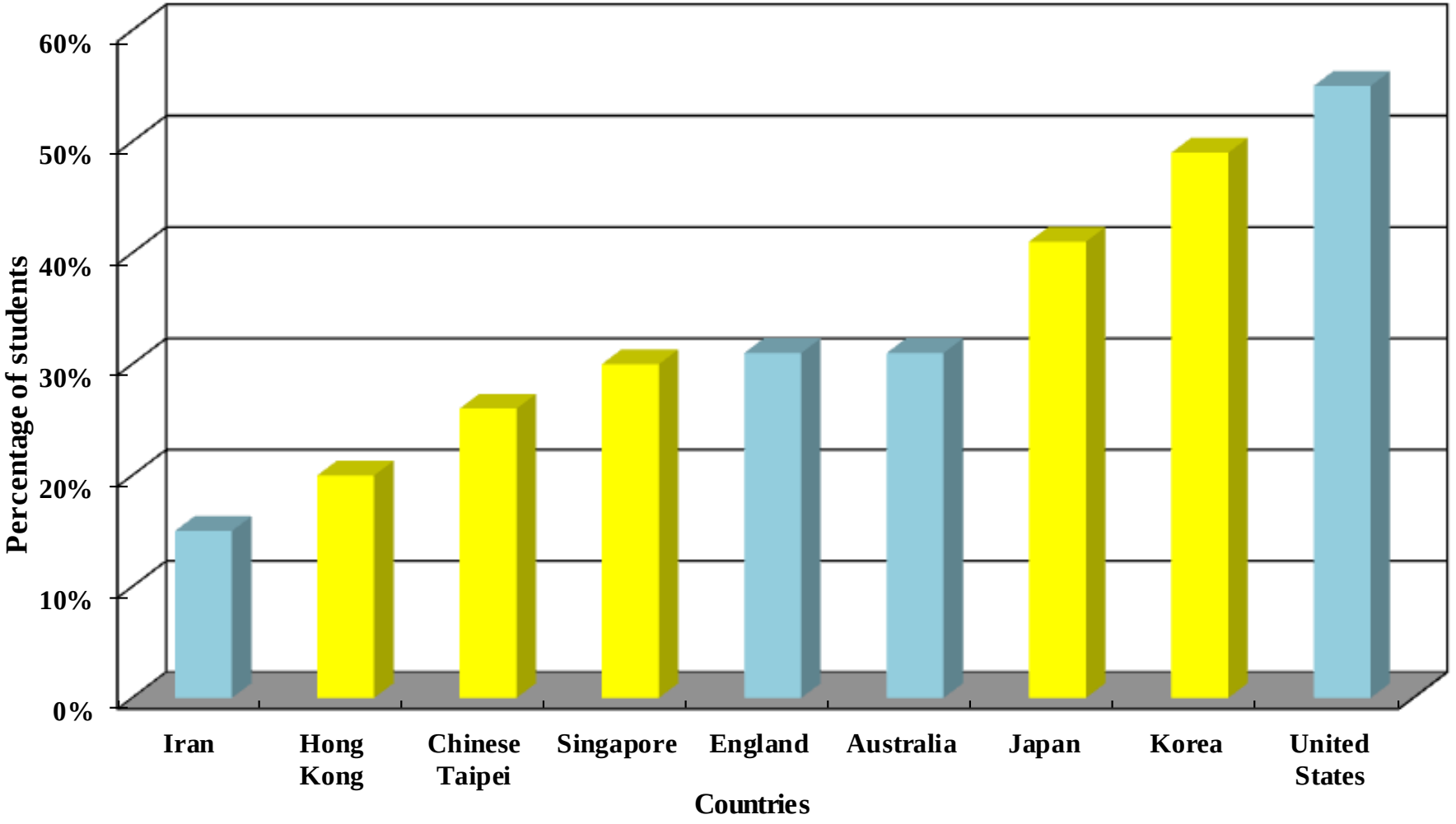
Secondary Pupil-Teacher Ratio



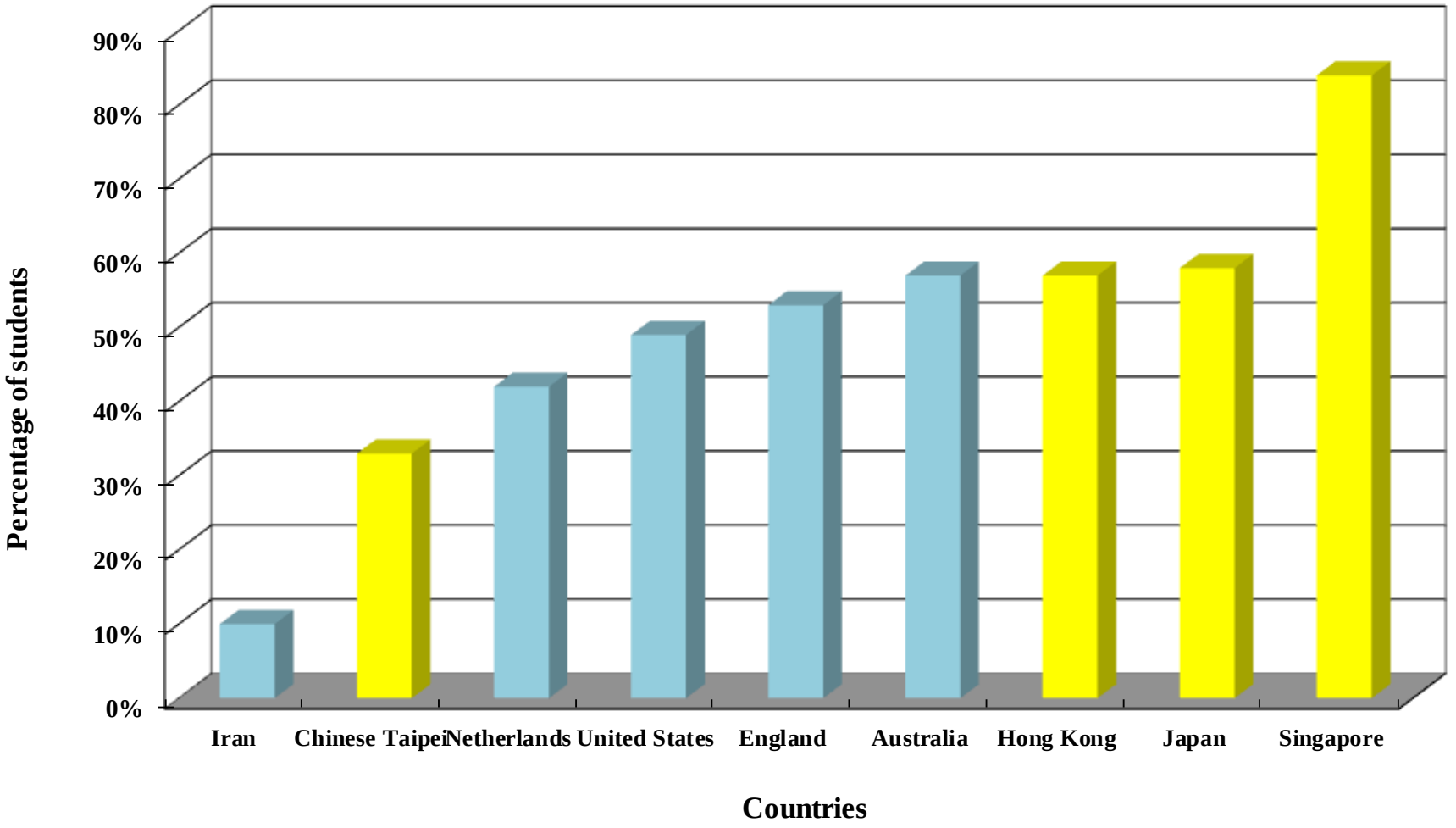
Weekly Time Students Spend on Assigned Mathematics Homework (G.8 Int'l Avg. = 15%)



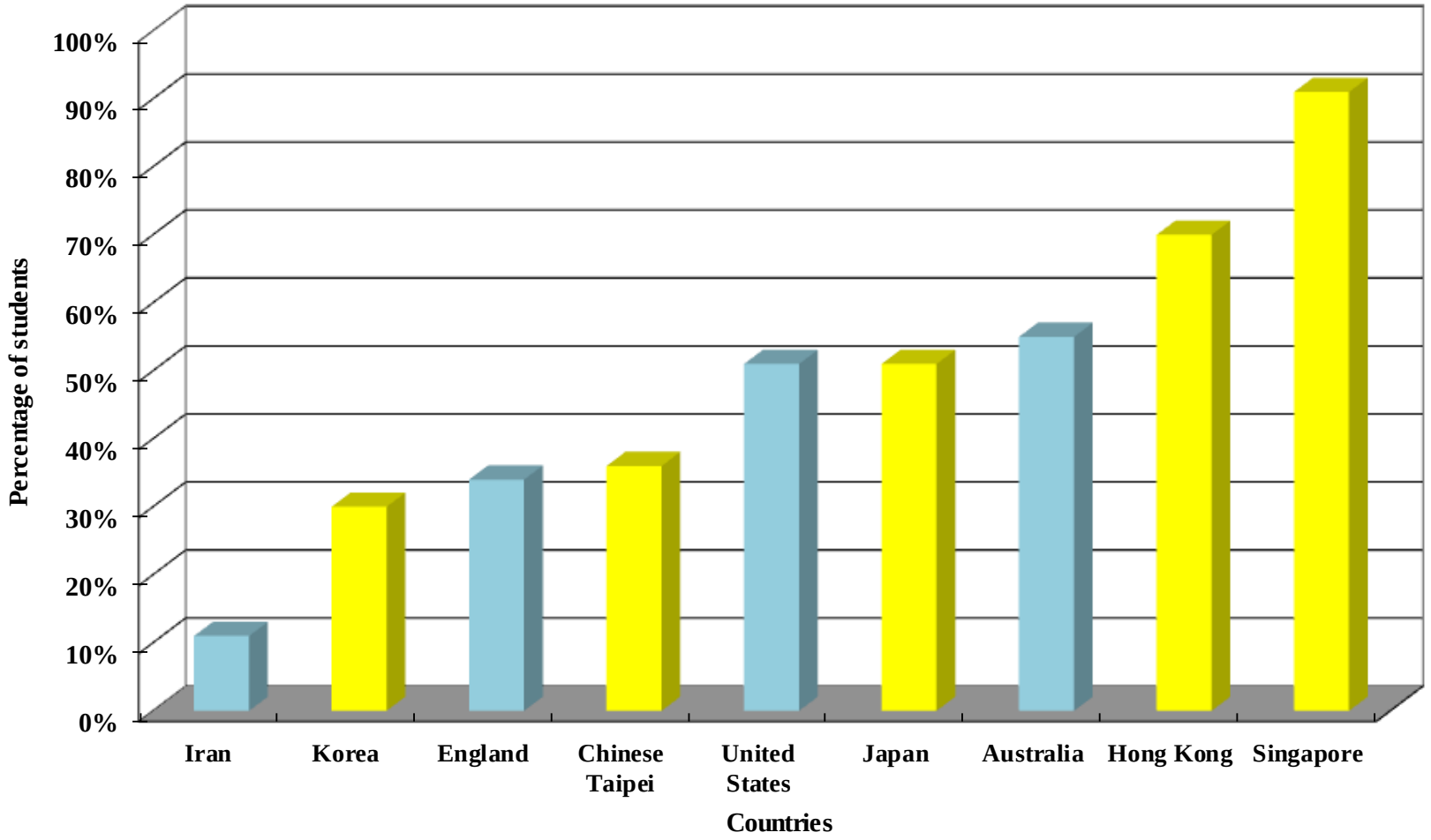
Parent with University Degree or Above (G8 Int'l Avg. = 32%)



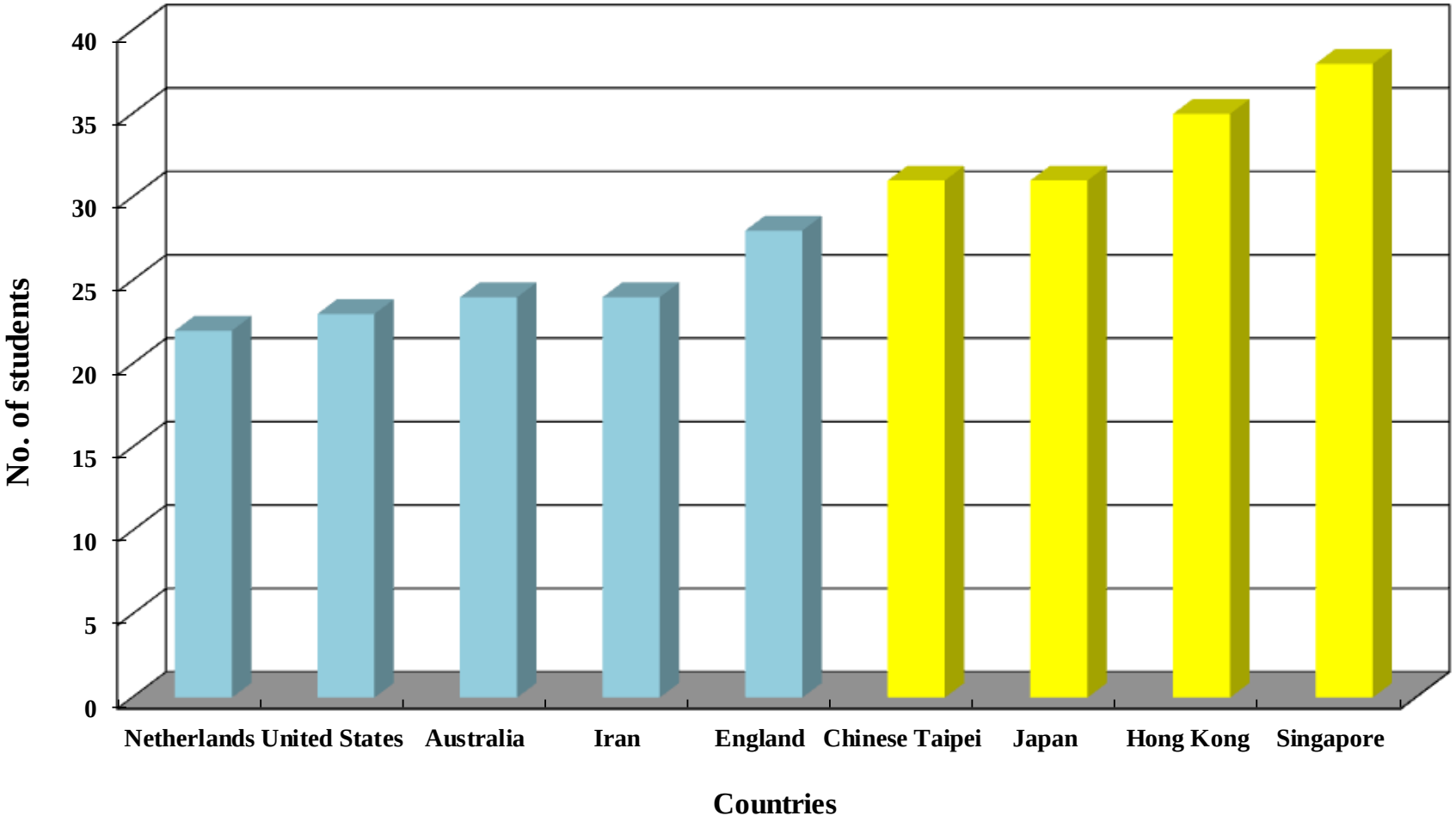
School Resources for Mathematics Instruction (G.4 Int'l Avg. = 43%)



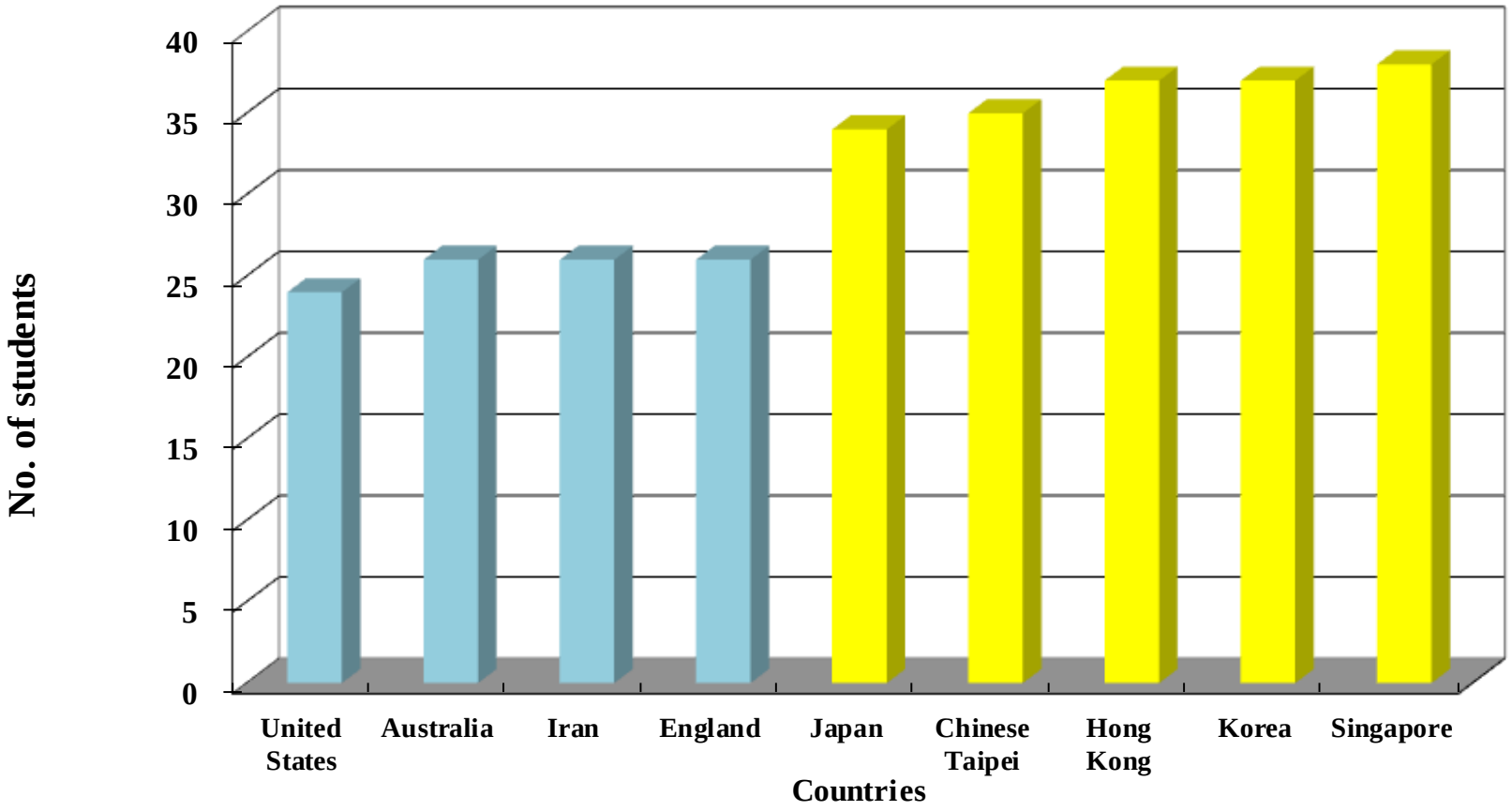
School Resources for Mathematics Instruction (G.8 Int'l Avg. = 32%)



Class size for Mathematics Instruction (G.4 Int'l Avg. = 26 students)



Class size for Mathematics Instruction (G.8 Int'l Avg. = 29 students)



What can we learn from these results?

No obvious relationship exists between these variables and student achievement, e.g.,

- Although Singapore, Japan and Hong Kong are relatively affluent as measured by GNI per capita, Taiwan and Korea are below average in wealth among the TIMSS countries
- Student/teacher ratio is relatively not favourable in Singapore, Korea and Japan
- Singapore and Hong Kong students spend a lot of time on homework, but this is not the case for Japan and Korea
- Schools in Singapore, Hong Kong and Japan are well resourced, but not those in Taiwan and Korea
- Class size in all five countries is large compared to Western countries

What can we conclude?

- Student achievement cannot be accounted for totally, or even to a significant extent, by these background factors
- Some factors (e.g., SES) may explain variations in achievement within a country, but they may not account for across country differences
- Background characteristics may be important factors for explaining student achievement, but they are not the only factors – and may not even be the most important factors

Culture as explanation for student achievement

- The five countries share a common culture, namely the Confucian heritage culture or CHC (Biggs, 1986)
- The underlying cultural values may be an important factor in explaining student achievement
- Low achievement does not necessarily imply the need of total revamp of the curriculum or instructional practices
- Complicated cultural factors might have affected classroom practices and student achievement, and so drastic changes should not be undertaken until such factors are thoroughly examined
- Any changes in educational policy must ensure that the strengths in a country are not lost in the process
- Simple transplant of policies and practices from high achieving countries to low achieving ones would not work, because one cannot transplant the practices without regard to the cultural differences

Which policy matters?
Which factors impact achievement?

E.g., Class size and student achievement

- Research question: does class size contribute to student achievement?
- It is extremely difficult for this research question to be answered by an educational experiment – random assignment of students to “experimental” and “control” group
- Research question best answered by these international studies
- What do the results tell us?
- Use TIMSS 2007 results as an example

Exhibit 7.2 **Achievement and Class Size for Mathematics Instruction**

TIMSS2007 Mathematics 4th Grade

Table 4

Country		1–19 Students		20–32 Students		33 or More Students	
		Percent of Students	Average Achievement	Percent of Students	Average Achievement	Percent of Students	Average Achievement
Algeria	r	11 (2.8)	388 (14.2)	60 (4.3)	378 (7.0)	29 (4.0)	383 (9.4)
Armenia	s	24 (3.3)	526 (14.1)	50 (3.8)	499 (7.3)	26 (3.6)	484 (6.0)
Australia		19 (3.0)	510 (9.0)	80 (3.0)	521 (4.3)	2 (1.2)	~ ~
Austria		37 (2.9)	506 (3.1)	63 (2.9)	505 (2.7)	0 (0.0)	~ ~
Chinese Taipei		3 (1.2)	548 (12.8)	45 (3.7)	570 (3.2)	51 (3.4)	583 (2.4)
Colombia		19 (3.3)	342 (13.7)	24 (4.7)	347 (14.0)	57 (4.4)	365 (8.1)
Czech Republic		31 (3.5)	482 (5.9)	69 (3.5)	489 (2.9)	0 (0.0)	~ ~
Denmark		34 (3.9)	529 (4.4)	66 (3.9)	521 (2.9)	0 (0.0)	~ ~
El Salvador		20 (2.7)	307 (10.7)	37 (4.1)	318 (9.1)	43 (3.8)	352 (4.2)
England		8 (1.9)	556 (9.6)	80 (3.0)	539 (3.2)	12 (2.4)	546 (9.0)
Georgia		37 (3.8)	454 (7.3)	50 (4.5)	428 (6.6)	13 (2.2)	454 (6.3)
Germany		21 (2.4)	512 (5.6)	79 (2.4)	528 (2.2)	0 (0.0)	~ ~
Hong Kong SAR		1 (0.7)	~ ~	25 (3.3)	588 (5.5)	74 (3.4)	616 (3.8)
Hungary		33 (3.7)	482 (6.5)	67 (3.7)	525 (4.7)	0 (0.0)	~ ~
Iran, Islamic Rep. of		25 (2.7)	381 (6.5)	59 (3.8)	406 (5.3)	16 (2.9)	421 (11.6)
Italy		44 (2.6)	506 (4.3)	56 (2.6)	507 (4.5)	0 (0.0)	~ ~
Japan		7 (1.5)	558 (8.5)	47 (2.9)	569 (3.4)	45 (3.2)	569 (2.9)
Kazakhstan		30 (4.5)	550 (20.2)	68 (4.6)	548 (5.5)	3 (1.2)	577 (29.4)
Kuwait	s	7 (2.8)	330 (18.1)	88 (3.4)	314 (5.0)	5 (1.9)	302 (11.9)
Latvia		44 (2.4)	525 (3.9)	49 (3.0)	550 (2.6)	6 (2.0)	551 (9.3)
Lithuania		37 (3.0)	511 (4.7)	63 (3.0)	541 (3.1)	0 (0.0)	~ ~
Morocco	r	17 (3.3)	352 (17.7)	42 (4.3)	343 (11.4)	41 (3.9)	338 (7.7)
Netherlands		27 (3.3)	531 (4.3)	71 (3.5)	535 (2.9)	2 (1.3)	~ ~
New Zealand	s	13 (2.1)	489 (8.7)	81 (2.4)	497 (3.0)	6 (1.7)	524 (11.7)
Norway		42 (3.3)	473 (4.4)	53 (3.6)	474 (3.5)	5 (1.9)	467 (10.6)
Qatar	r	8 (0.1)	301 (4.3)	75 (0.2)	296 (1.4)	17 (0.2)	316 (3.4)
Russian Federation		33 (2.7)	531 (10.5)	67 (2.7)	551 (3.8)	0 (0.3)	~ ~
Scotland	r	16 (2.8)	492 (9.4)	79 (3.0)	493 (3.1)	5 (1.6)	506 (14.0)
Singapore		0 (0.0)	~ ~	6 (1.3)	514 (13.5)	94 (1.3)	605 (3.5)
Slovak Republic		34 (2.5)	497 (6.6)	65 (2.6)	496 (5.7)	1 (0.6)	~ ~
Slovenia		46 (2.9)	497 (2.7)	53 (3.0)	506 (2.6)	1 (0.6)	~ ~
Sweden		36 (3.4)	505 (4.5)	60 (3.6)	504 (3.2)	4 (1.6)	512 (12.4)
Tunisia		20 (2.8)	303 (12.2)	69 (3.8)	334 (5.0)	11 (2.7)	354 (21.3)
Ukraine		30 (3.3)	445 (4.9)	65 (3.5)	480 (3.8)	5 (1.4)	472 (13.4)
United States		26 (2.6)	521 (4.1)	69 (2.8)	533 (3.3)	5 (1.3)	522 (8.0)
Yemen	r	9 (2.1)	262 (18.5)	17 (4.0)	227 (16.4)	74 (4.1)	219 (7.7)
International Avg.		24 (0.5)	462 (1.8)	58 (0.6)	471 (1.1)	18 (0.4)	460 (2.3)

SOURCE: IEA's Trends in International Mathematics and Science Study (TIMSS) 2007

t and Class Size for Mathematics Instruction

	1–19 Students		20–32 Students		3
	Percent of Students	Average Achievement	Percent of Students	Average Achievement	Per of St
r	11 (2.8)	388 (14.2)	60 (4.3)	378 (7.0)	29
s	24 (3.3)	526 (14.1)	50 (3.8)	499 (7.3)	26
	19 (3.0)	510 (9.0)	80 (3.0)	521 (4.3)	2
	37 (2.9)	506 (3.1)	63 (2.9)	505 (2.7)	0
	3 (1.2)	548 (12.8)	45 (3.7)	570 (3.2)	51
	19 (3.3)	342 (13.7)	24 (4.7)	347 (14.0)	57
	31 (3.5)	482 (5.9)	69 (3.5)	489 (2.9)	0
	34 (3.9)	529 (4.4)	66 (3.9)	521 (2.9)	0
	20 (2.7)	307 (10.7)	37 (4.1)	318 (9.1)	43
	8 (1.9)	556 (9.6)	80 (3.0)	539 (3.2)	12
	37 (3.8)	454 (7.3)	50 (4.5)	428 (6.6)	13
	21 (2.4)	512 (5.6)	79 (2.4)	528 (2.2)	0
	1 (0.7)	~ ~	25 (3.3)	588 (5.5)	74
	33 (3.7)	482 (6.5)	67 (3.7)	525 (4.7)	0
. of	25 (2.7)	381 (6.5)	59 (3.8)	406 (5.3)	16
	44 (2.6)	506 (4.3)	56 (2.6)	507 (4.5)	0
	7 (1.5)	558 (8.5)	47 (2.9)	569 (3.4)	45
	30 (4.5)	550 (20.2)	68 (4.6)	548 (5.5)	3
s	7 (2.8)	330 (18.1)	88 (3.4)	314 (5.0)	5
	44 (2.4)	525 (3.9)	49 (3.0)	550 (2.6)	6
	37 (3.0)	511 (4.7)	63 (3.0)	541 (3.1)	0
r	17 (3.3)	352 (17.7)	42 (4.3)	343 (11.4)	41

Has a relation been established between class size and student achievement according to the data?

- For many countries (e.g., Austria, Italy), class size does not make any difference to student achievement
- For some countries (e.g., Armenia, Kuwait), the smaller the class size, the higher the student achievement
- For the majority of the countries (e.g., Chinese Taipei, Colombia, New Zealand), the bigger the class size, the higher the student achievement
- All the high achieving countries (e.g., Singapore, Korea, Hong Kong) have large class sizes
- How do these results guide “educational decision making and practice”?
- Are we going to suggest increasing class size in order to raise the achievement of students??

Limitations of large scale international studies

Comparability problems:

- Sample: grade or age? What is grade 8? Is comparing 15 year olds around the world "fair"?
- System differences: e.g., application of decimals in currencies problems (the use of "zed" in TIMSS)
- Language
  - Equivalence in the translation of instruments (TIMSS and PISA involve more than 60 countries operating in more than 30 languages; some items become meaningless after translation (e.g., "How many sides are there in a heptagon?"))
  - Does language affect the way we process mathematics in the test matter?

The root of the problem

- In international studies, we compare across cultures, using the world as “a natural educational laboratory”
- Many variables within a country or culture are uniform and cannot be manipulated, and to study the impact of those variables on student achievement, we have to collect data in different cultures, where the variables differ
- But not only are those variables of interest differ, a host of other variables are vastly different as well, and usually these variables exist as a bundle
- So it is difficult, if not impossible, to control for all the other variables in studying the variables of interest
- And we are never sure whether we have taken all relevant variables into account
- Husen (1983): in international studies, “**we are comparing the incomparables**”!

Is it legitimate to rank countries?

- Rigorous methodology adopted in these studies means results on student achievement rather reliable
- So methodologically speaking, the data of these studies do allow us to rank countries
- But we need to be careful in interpreting rankings
- Participating countries in these studies change from one cycle to another, so a rank of say 20th in a certain cycle may not mean the same thing as a rank of 20th in another cycle
- Also, when comparing the relatively rankings between two countries, we should take the standard error of measurement into consideration

e.g., Singapore TIMSS 2003 and 2007

- Compared to TIMSS 2003, grade 8 students in Singapore may be seen as "dropping" from the first place to the third place in TIMSS 2007
- But if we take the standard errors of measurement into consideration, the differences between the score for Singapore and those of Korea (rank 2) and Chinese Taipei (rank 1) in 2007 are not statistically significant
- From a statistical point of view, we cannot say that the scores of Chinese Taipei and Korea are higher than that of Singapore
- So we should not be too sensitive about fine changes in ranking from cycle to cycle - it is usually not too meaningful to say that a country's ranking has dropped from say 15^{th} to 18^{th} without further qualification

Exhibit 1.1 TIMSS 2007 Distribution of Mathematics Achievement (Continued)

TIMSS2007 Mathematics 8th Grade

Country	Mathematics Achievement Distribution	Average Scale Score	Years of Formal Schooling*	Average Age at Time of Testing	Human Development Index**
Chinese Taipei		▲ 598 (4.5)	8	14.2	0.932
Korea, Rep. of		▲ 597 (2.7)	8	14.3	0.921
Singapore		▲ 593 (3.8)	8	14.4	0.922
† Hong Kong SAR		▲ 572 (5.8)	8	14.4	0.937
Japan		▲ 570 (2.4)	8	14.5	0.953
Hungary		▲ 517 (3.5)	8	14.6	0.874
† England		▲ 513 (4.8)	9	14.2	0.946
Russian Federation		▲ 512 (4.1)	7 or 8	14.6	0.802
2 † United States		▲ 508 (2.8)	8	14.3	0.951
1 Lithuania		▲ 506 (2.3)	8	14.9	0.862
Czech Republic		504 (2.4)	8	14.4	0.891
Slovenia		501 (2.1)	7 or 8	13.8	0.917
TIMSS Scale Avg.		**500**			
Armenia		499 (3.5)	8	14.9	0.775
Australia		496 (3.9)	8	13.9	0.962
Sweden		▼ 491 (2.3)	8	14.8	0.956
Malta		▼ 488 (1.2)	9	14.0	0.878
† Scotland		▼ 487 (3.7)	9	13.7	0.946
1 2 Serbia		▼ 486 (3.3)	8	14.9	0.810
Italy		▼ 480 (3.0)	8	13.9	0.941
Malaysia		▼ 474 (5.0)	8	14.3	0.811
Norway		▼ 469 (2.0)	8	13.8	0.968
Cyprus		▼ 465 (1.6)	8	13.8	0.903

SOURCE: IEA's Trends in International Mathematics and Science Study (TIMSS) 2007

Is it fair to rank countries?

Test-curriculum match

- Given the scale of these studies, the test items inevitably match the curricula in some countries better than others
- This problem is particularly acute in TIMSS as it aims at testing students' competence according to the school curriculum
- e.g., while there is 100% coverage of the TIMSS 2015 grade 4 test items in Saudi Arabia, the coverage for the Slovak Republic curriculum is only 57.3% and for Lebanon 68.8%
- (Though the test-curriculum match analysis shows that curricular differences do not make significant difference to countries' average scores and rankings)

Can we draw causal relations?

- These international studies are surveys, and not experiments
- So we have to be extra cautious in drawing conclusions about causal relations
- In most instances, the best that we can conclude is that a certain variable A ***may*** have caused or impacted student achievement, based on the correlations between the measure of variable A and the achievement scores, since it is unlikely or illogical that achievement leads to changes in variable A
- But we cannot rule out the possibility that there is a third "hidden" variable which influences both variable A and achievement, causing variable A and achievement to be correlated with each other
- And there are so many possible variables that may have influenced both variable A and achievement!

Examples

(1) Class size and achievement

Does big class size lead to high achievement, or are there variables which lead to both large class size and high achievement?

(2) The relation between amount of homework and achievement

Students may have better achievement because they do more homework, but students may need to do more homework because they have low achievement

It is therefore not surprising that there is no clear relationship between student achievement and the amount of homework students do.

Can we compare teachers in different countries?

- If we need to be careful in drawing any conclusions about student attributes (achievement, attitudes, etc.), we need to be even more careful in drawing conclusions about teacher attributes and performances in different countries
- This is because in TIMSS, we do not have a representative sample of teachers, so all references to teachers are from the student sample
- For example, according to the TIMSS 2015 findings, we cannot say that 93% of the primary school teachers in Denmark have a university degree or above, all we can only say is that 93% of the primary (four) school students have teachers with a university degree or above.
- Any policy suggestions about teachers should bear this in mind

Policy implications of these studies

- Despite the aims of these international studies in providing information for "guiding educational decision making and practice in the areas of mathematics (and science)", we should be extremely cautious in suggesting changes in policy and practice based results of these studies, given the limitations of these studies and the dubious nature of drawing causal relations from the findings
- Actually in many instances, curriculum changes claimed to be based on results of these international studies were more excuses rather than rational decisions based on a rigorous examination of the results.

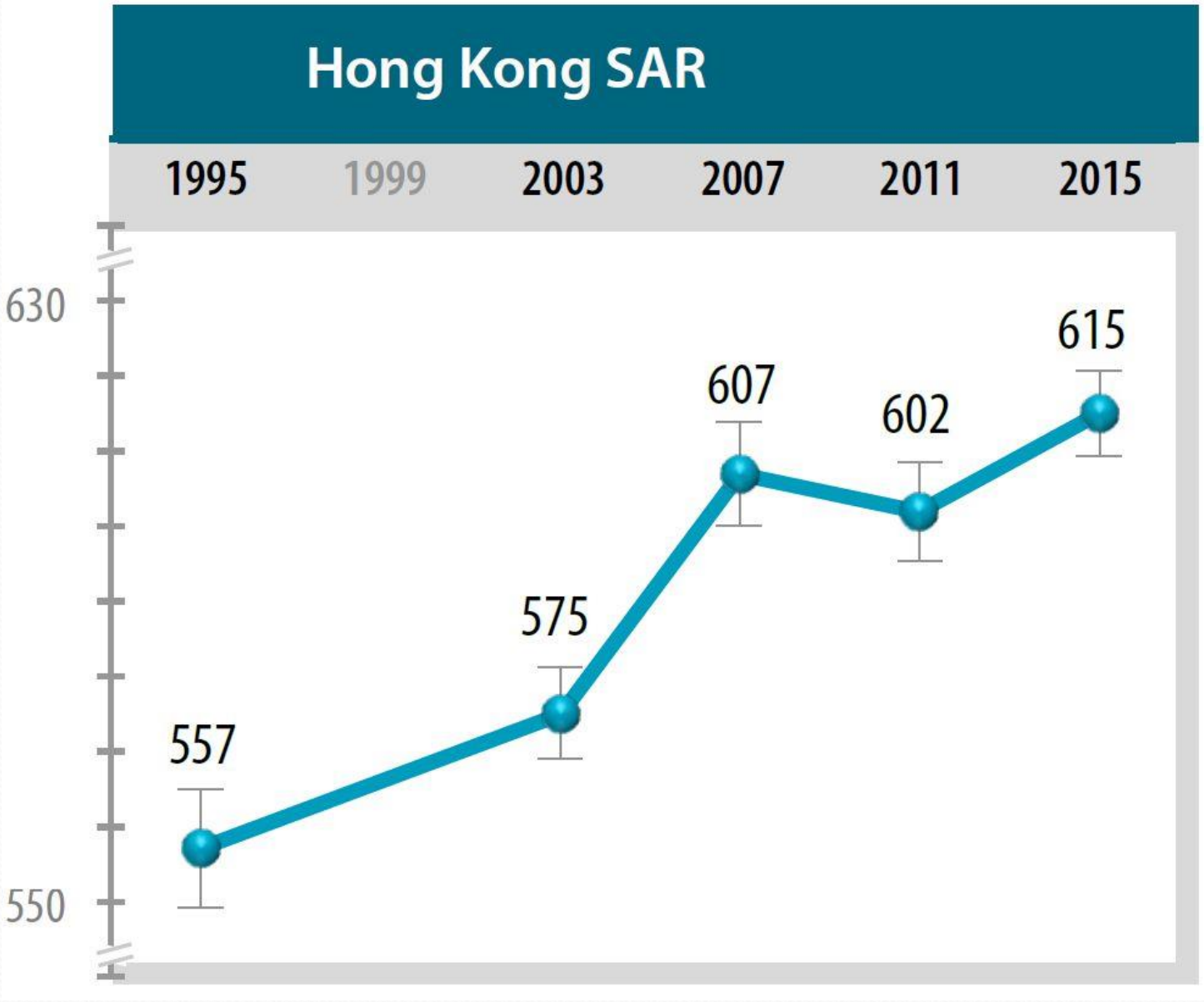
What can we learn from these studies?

- Despite all the limitations of these international studies, the rigorous methodologies adopted in these studies do provide us with a reliable measure of student achievement, and hence “effectiveness” of an education system
- Since these studies are “international (studies) with endorsement from a large number of countries”, they provide benchmarks against which countries may measure the achievement of their students
- What can we learn from these studies?

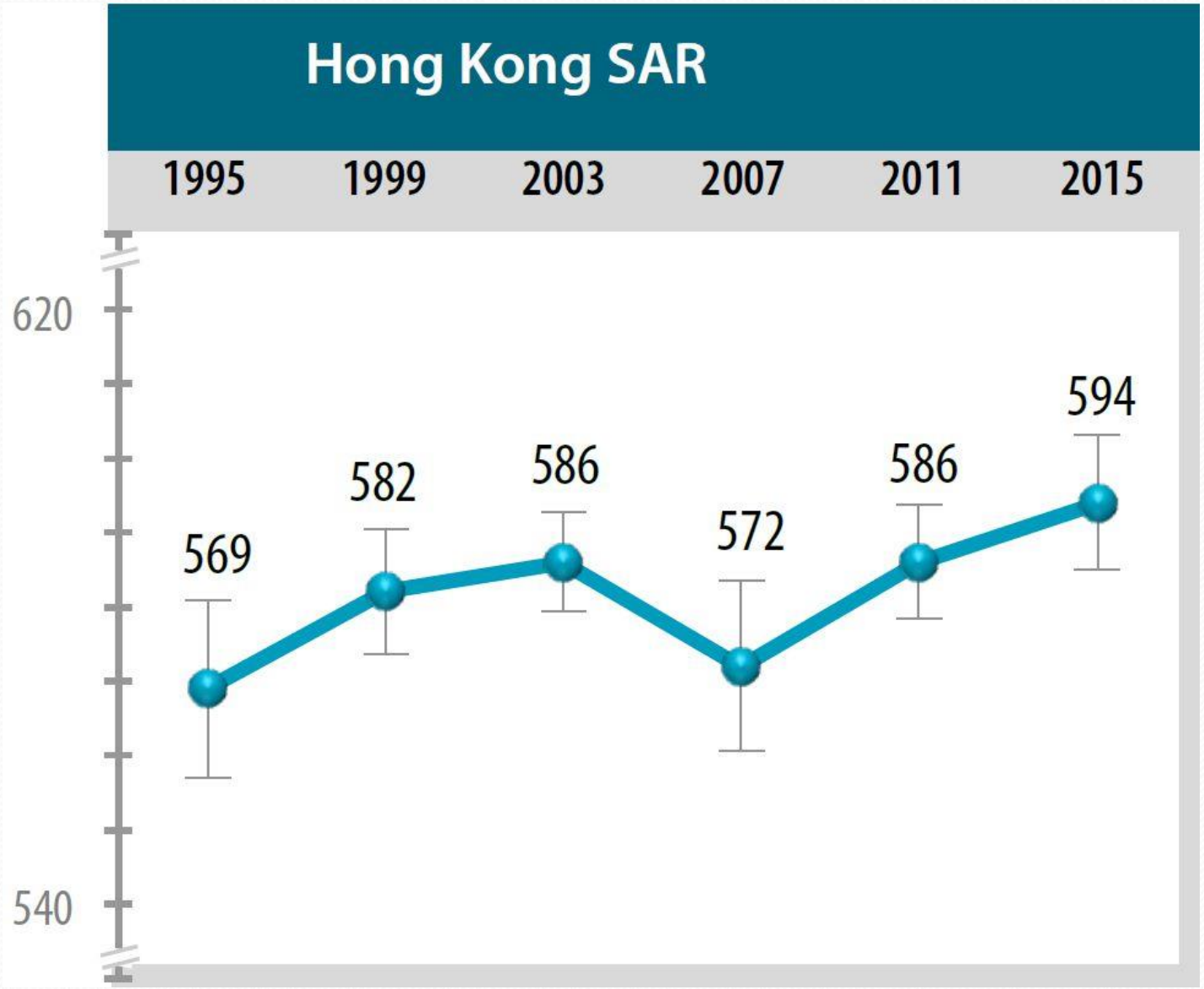
1. Trend of student achievements

- For those countries which have participated in more than one cycle of the studies, a very instructive piece of information is the change of scores (rather than change of ranking) across different cycles
- Scores in these studies are standardized across years and are thus theoretically comparable
- But these are not truly longitudinal studies
- E.g., when the scores of TIMSS 2011 grade 4 students in a certain country are compared to the TIMSS 2015 grade 8 students, the students come from the same cohort but not the same students were taking the tests, so any “gain” in scores only gives rough indication of “trends”
- Notwithstanding this limitation, this rough information on trends of performance should be informative to educators in the country, especially when there are major curriculum changes taking place in between the cycles of study
- Look at the trends in mathematics achievements in TIMSS for Hong Kong

Grade 4


Hong Kong SAR
1995
1999
2003
2007
2011
2015
630
550
557
575
607
602
615

Grade 8


Hong Kong SAR
1995
1999
2003
2007
2011
2015
620
540
569
582
586
572
586
594

2. Comparison of performance in different strands and areas of the curriculum

- Performance in different strands of mathematics (content strand, e.g., geometry versus statistics; cognitive domain, e.g., reasoning versus knowing), will inform us of the relative strengths and weaknesses of our students in light of the performance of students in another country or internationally
- Hong Kong students are not doing too well in the domain of “Data Display” and in “Reasoning”
- As Statistics is becoming more important in the contemporary world, it is important to lay a solid foundation on the basic concepts of Statistics from the early stages
- In this modern age when generic skills are much more important than mastery of specific knowledge and skills, perhaps more weight should be given to developing the reasoning abilities in students

Performance of Hong Kong students

Primary 4	Number	Geometric Shapes and Measures	Data Display
HKSAR	616	617	611

	Knowing	Applying	Reasoning
HKSAR	618	621	600

Secondary 2	Number	Algebra	Geometry	Data and Chance
HKSAR	594	593	602	597

	Knowing	Applying	Reasoning
HKSAR	600	595	591

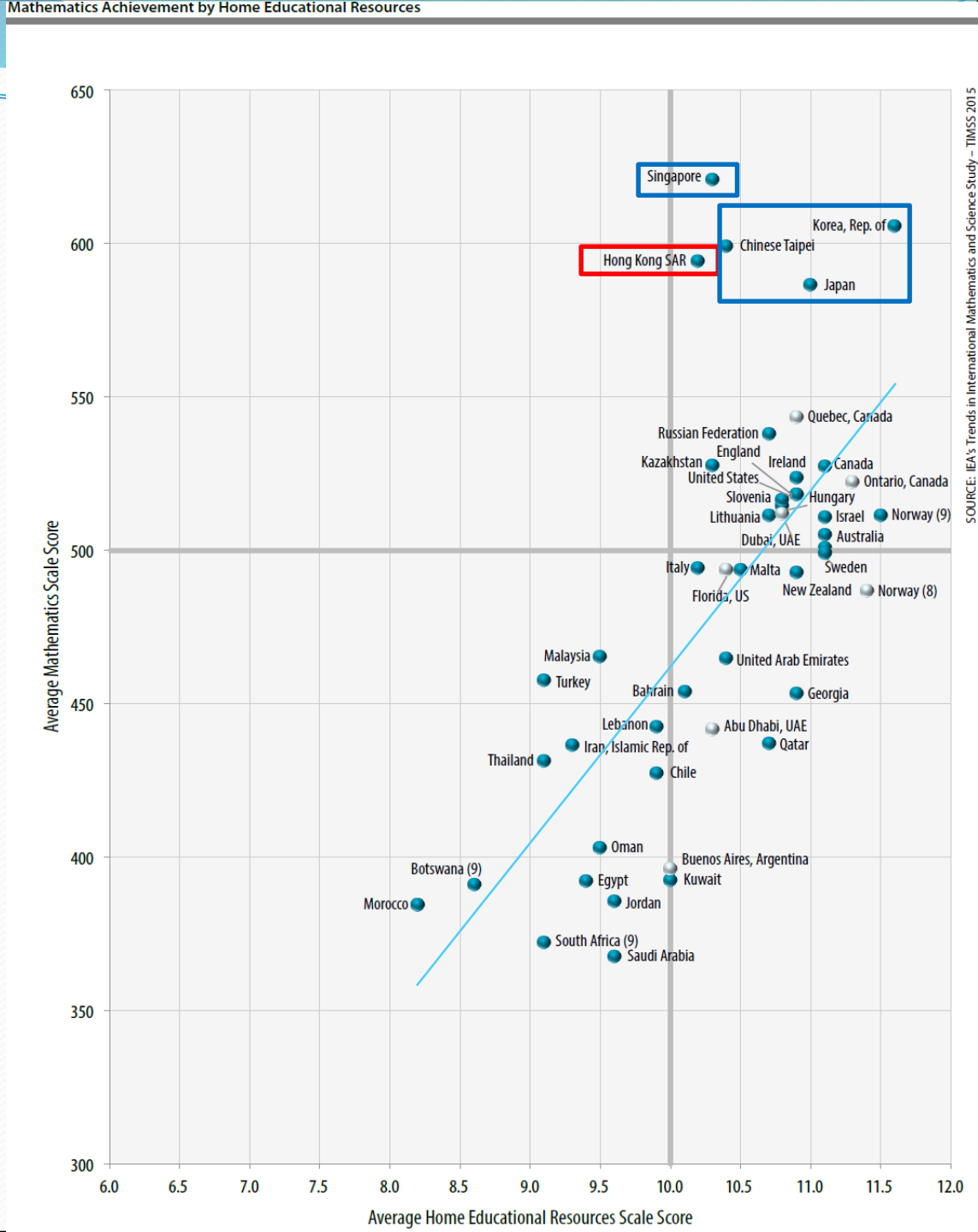
3. Effectiveness of the system

Grade 4

Home Resources for Learning

Grade 8

Home Educational Resources

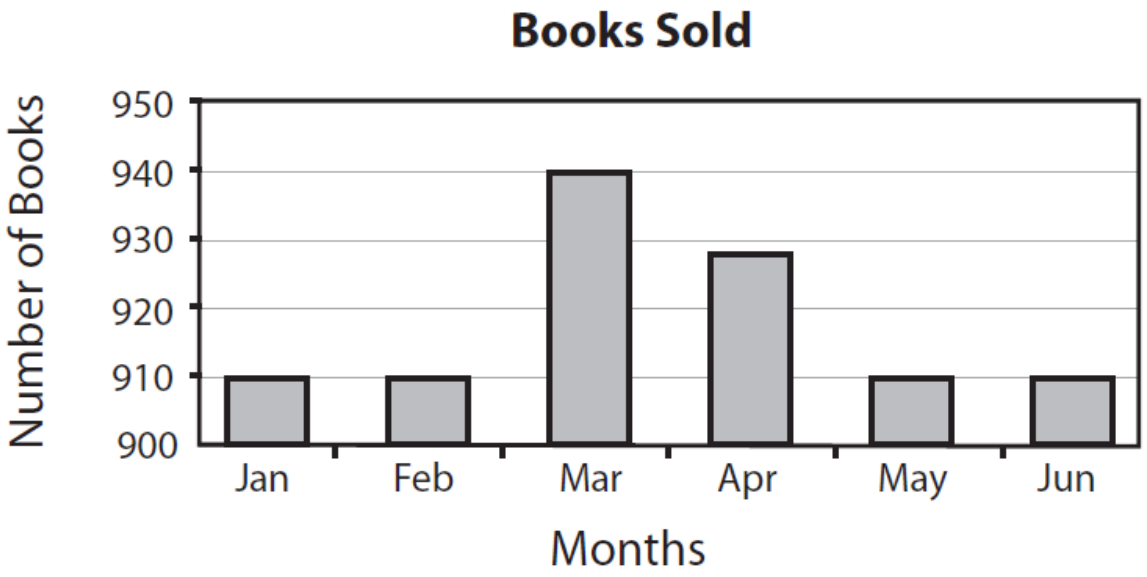


4. What can teachers learn from these studies: Two-digit diagnostic codes

- In the scoring of open-ended items of the TIMSS test, a two-digit scoring code is used, the first digit records the marks given to that item (partial correct answers are reflected by the marks awarded), while the second digit categories how the student arrives at the right or wrong answer
- The second digit will inform us of the typical way the item is solved in a country or a school, and more importantly typical misconceptions concerning that item
- These are extremely useful information for teachers

Two-digit Diagnostic Codes

- Example: M01_14 (Data and Chance / Reasoning)



A salesman looked at the graph showing his sales of books for the first 6 months of 2004, and said, "In March, I sold four times as many books as I sold in February."

Explain whether you agree or disagree with the salesman, and give a reason.

Note: Code 10 takes precedence over other correct explanation (code 11 and code 12).
Code 11 takes precedence over code 12.

Code	Response	Item: M042164
	Correct Response	
10	Disagree, with reference to false origin or scale not starting from zero *Examples:* *I disagree because the graph section of the number of books does not start at zero.* *I disagree with the salesman. He should look at the graph carefully. The graph is plotted using 900 as the base and not 0.*	
11	Disagree with explanation based on multiplication or division *Examples:* *I disagree because I do not think that 940 is 4 times as many. I think if it is 4 times as many it would be 3640.* *Disagree. As the graph shows that he sold 910 books in February and 940 books in March. 940 is not 4 times of 910.* *I disagree because if you divide March's total by 4 (940Z ÷ 4), you get 235, which is not February's total. February's total was 910*	
12	Disagree, with explanation that the increase cannot be 4 times as many books. *Example:* *I disagree because he only sold 30 more books in March. From 910 books he went to 940.*	
	Incorrect Response	
70	Agree or disagree, with explanation based on only relative heights of the bars shown *Examples:* *I agree because the graph shows that in March the bar went up four times.* *I disagree as if you look at the bar you will see between February and March there are only three bars, so he would have to say, "In March I sold three times as many books as I did in February."*	
79	Other incorrect (including crossed out, erased, stray marks, illegible, or off task) *Examples:* *I agree because in February he sold 30 less book then March.* *I disagree with the salesman because in February he sold 910 books. That would mean he would have sold 1820 books but he only sold 940 books in March.*	
	Nonresponse	
99	Blank	

Two-digit Diagnostic Codes

- Example: M01_14 (Data and Chance / Reasoning)

	10	11	12	70	79	V1	OMITTED	NOT REACHED	GIRLS	BOYS
HONG KONG	8.8	53.4	11.2	1.2	20.6	73.5	3.7	1.0	71.9	74.9
CHINESE TAIPEI	2.4	53.8	7.5	2.2	29.0	63.6	4.2	1.0	66.4	60.9
JAPAN	6.1	37.1	14.7	10.1	22.3	57.8	8.4	1.4	59.4	56.2
KOREA	2.2	52.9	15.8	7.0	17.5	71.0	3.7	0.8	72.0	70.1
SINGAPORE	8.3	58.5	7.2	5.6	16.4	74.0	2.3	1.7	78.4	69.9
INT'L AVG	3.1	23.8	8.2	7.3	40.1	35.1	12.5	5.0	35.4	34.8

Two-digit Diagnostic Codes

- Example: M01_02 (Number / Knowing)

Write this as a decimal number.

$$8 + 50 + \frac{3}{100} + \frac{1}{10}$$

Answer: _____

Two-digit Diagnostic Codes

- Example: M01_02 (Number / Knowing)

ID: **M042081**	Mathematics Grade 8	Block_Seq: M01_02

Code	Response	Item: M042081
	Correct Response	
10	58.13	
	Incorrect Response	
70	$58\frac{13}{100}$	
71	58.31	
72	85.31	
79	Other incorrect (including crossed out, erased, stray marks, illegible, or off task)	
	Nonresponse	
99	Blank	

Two-digit Diagnostic Codes

- Example: M01_02 (Number / Knowing)

	10	70	71	72	79	V1	OMITTED	NOT REACHED	GIRLS	BOYS
HONG KONG	68.9	8.4	0.0	0.0	19.6	68.9	3.1	0.0	70.0	67.9
CHINESE TAIPEI	**54.0**	2.0	0.5	0.0	37.9	54.0	5.6	0.0	52.2	55.8
JAPAN	**48.4**	0.9	0.0	0.0	40.8	48.4	9.8	0.1	49.2	47.5
KOREA	**61.3**	0.7	0.0	0.1	34.1	61.3	3.6	0.2	60.2	62.3
SINGAPORE	**85.7**	2.6	0.0	0.0	10.5	85.7	1.1	0.1	87.6	84.0
INT'L AVG	**35.4**	2.2	0.5	0.2	48.6	35.4	13.0	0.2	36.8	34.0

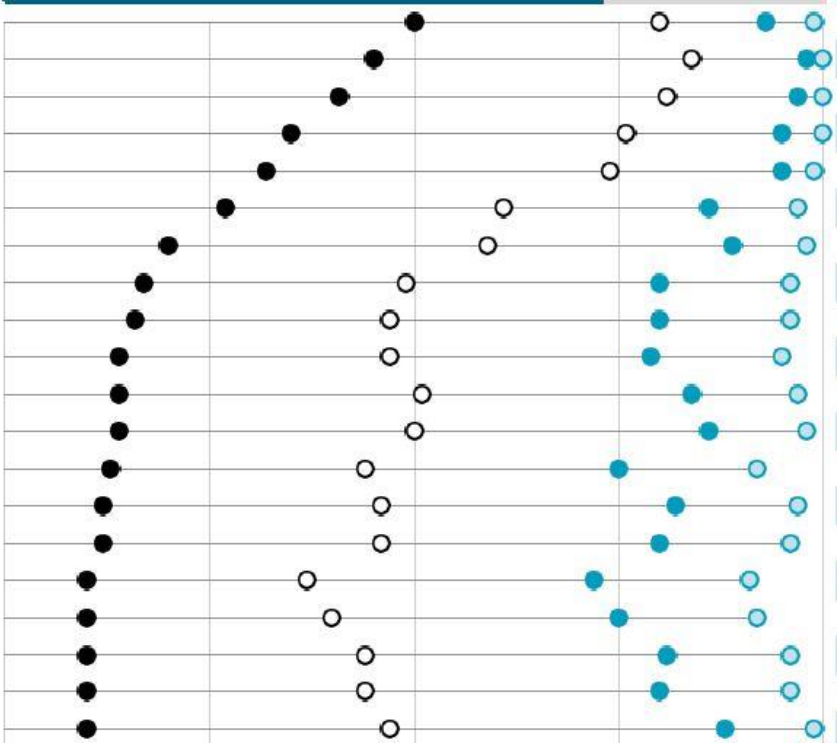
5. Policy implications: e.g., how equitable is the education provision in Hong Kong?

A. International benchmarks

- Nearly all Primary 4 Hong Kong students attain the Low International Benchmark in mathematics
- But there are 2% of the Hong Kong Secondary 2 students who fail to attain even this lowest benchmark
- There are substantial proportions of Hong Kong students attaining Advanced International Benchmark in mathematics
- But the proportion of such students in Secondary 2 is significantly lower than those in other high performing countries

International Benchmarks

Primary 4 (first 20 countries)

Country	Percentages of Students Reaching International Benchmarks (● Advanced ○ High ● Intermediate ○ Low)	Advanced Benchmark (625)	High Benchmark (550)	Intermediate Benchmark (475)	Low Benchmark (400)
[2] Singapore		50 (2.1)	80 (1.7)	93 (0.9)	99 (0.3)
† Hong Kong SAR		45 (2.0)	84 (1.3)	98 (0.4)	100 (0.1)
Korea, Rep. of		41 (1.3)	81 (1.0)	97 (0.4)	100 (0.1)
Chinese Taipei		35 (1.5)	76 (1.0)	95 (0.4)	100 (0.2)
Japan		32 (1.1)	74 (1.0)	95 (0.4)	99 (0.1)
‡ Northern Ireland		27 (1.3)	61 (1.5)	86 (1.1)	97 (0.6)
Russian Federation		20 (1.8)	59 (1.8)	89 (1.1)	98 (0.4)
England		17 (1.2)	49 (1.5)	80 (1.2)	96 (0.7)
Kazakhstan		16 (1.8)	47 (2.6)	80 (1.5)	96 (0.5)
[2] † United States		14 (0.8)	47 (1.1)	79 (1.0)	95 (0.5)
Ireland		14 (1.0)	51 (1.6)	84 (1.0)	97 (0.4)
Norway (5)		14 (1.1)	50 (1.6)	86 (1.0)	98 (0.4)
Hungary		13 (0.9)	44 (1.5)	75 (1.5)	92 (0.9)
[2] Portugal		12 (0.9)	46 (1.3)	82 (1.1)	97 (0.4)
[2] † Denmark		12 (0.9)	46 (1.6)	80 (1.3)	96 (0.6)
[3] Serbia		10 (0.8)	37 (1.4)	72 (1.6)	91 (1.2)
Bulgaria		10 (1.3)	40 (2.6)	75 (2.1)	92 (1.3)
[2] Lithuania		10 (1.0)	44 (1.5)	81 (1.1)	96 (0.5)
Poland		10 (0.7)	44 (1.4)	80 (1.0)	96 (0.4)
† Belgium (Flemish)		10 (0.8)	47 (1.5)	88 (0.9)	99 (0.3)
International Median		6	36	75	93

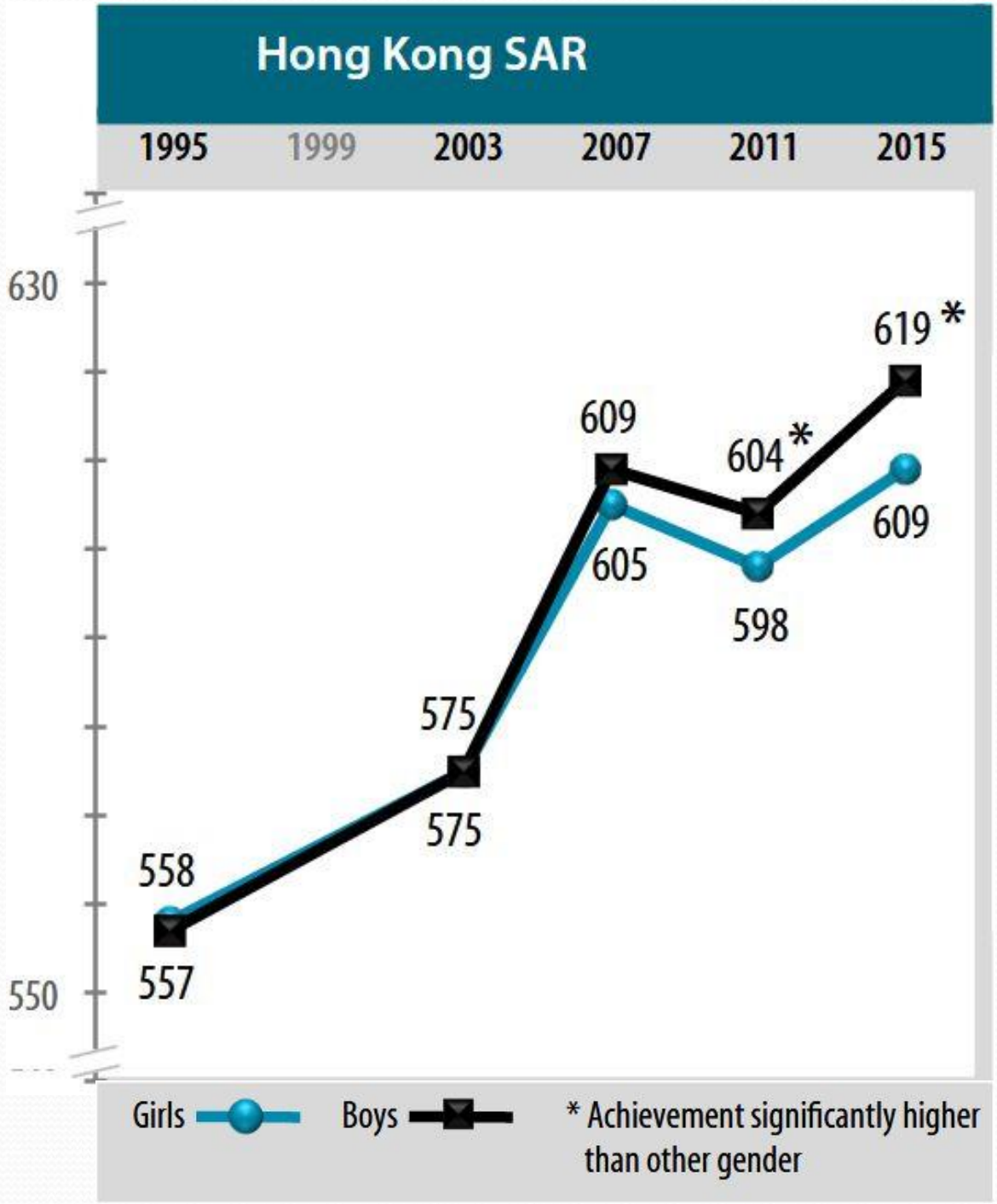
International Benchmarks

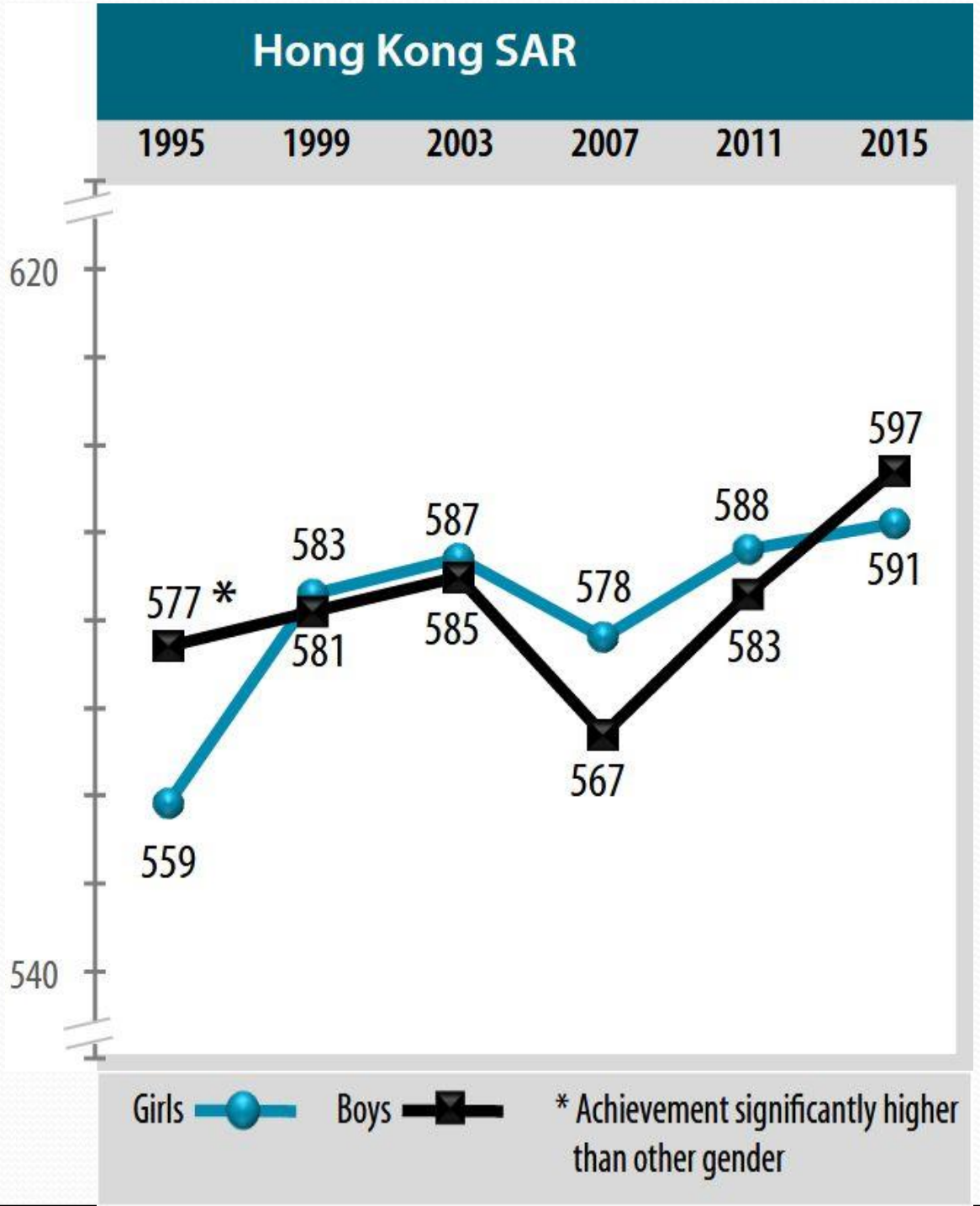
Secondary 2 (first 20 countries)

Country	Percentages of Students Reaching International Benchmarks (● Advanced ○ High ● Intermediate ○ Low)	Advanced Benchmark (625)	High Benchmark (550)	Intermediate Benchmark (475)	Low Benchmark (400)
[2] Singapore		54 (1.8)	81 (1.5)	94 (0.9)	99 (0.2)
Chinese Taipei		44 (1.2)	72 (0.9)	88 (0.6)	97 (0.4)
Korea, Rep. of		43 (1.4)	75 (1.0)	93 (0.5)	99 (0.2)
Hong Kong SAR		37 (2.3)	75 (1.9)	92 (1.3)	98 (0.6)
Japan		34 (1.2)	67 (1.0)	89 (0.7)	98 (0.3)
Kazakhstan		15 (1.7)	41 (2.6)	71 (2.1)	91 (1.1)
Russian Federation		14 (1.4)	46 (2.5)	78 (1.9)	95 (0.8)
[3] Israel		13 (1.0)	38 (1.8)	65 (1.7)	84 (1.2)
Hungary		12 (1.2)	37 (1.7)	67 (1.7)	88 (1.1)
† United States		10 (0.9)	37 (1.5)	70 (1.4)	91 (0.7)
England		10 (1.1)	36 (2.4)	69 (2.4)	93 (1.2)
[1] † Canada		7 (0.6)	39 (1.4)	78 (1.1)	96 (0.5)
Australia		7 (0.8)	30 (1.4)	64 (1.6)	89 (1.0)
Ireland		7 (0.8)	38 (1.7)	76 (1.3)	94 (0.8)
† New Zealand		6 (0.8)	27 (1.2)	58 (1.5)	85 (1.2)
Turkey		6 (0.9)	20 (1.6)	42 (1.9)	70 (1.6)
[2] Lithuania		6 (0.8)	33 (1.4)	68 (1.4)	92 (0.8)
Slovenia		6 (0.6)	32 (1.3)	73 (1.2)	95 (0.6)
United Arab Emirates		5 (0.4)	20 (0.8)	46 (1.0)	73 (0.7)
Malta		5 (0.4)	29 (0.7)	62 (0.7)	84 (0.5)
International Median		5	26	62	84

B. Gender differences

- Gender difference in achievement has not been a problem, but in 2011 and 2015, for the first time since Hong Kong participated in TIMSS, Primary 4 boys outperformed girls in mathematics


Hong Kong SAR
1995
1999
2003
2007
2011
2015
630
550
558
557
575
575
609
605
604 *
598
619 *
609
Girls
Boys
* Achievement significantly higher than other gender


Hong Kong SAR
1995
1999
2003
2007
2011
2015
620
540
577 *
581
583
585
587
567
578
583
588
591
597
559
Girls
Boys
* Achievement significantly higher than other gender

C. Socioeconomic status

- For socioeconomic status, there is a statistically significant relation between all the variables and the mathematics achievement of Primary 4 students
- Conclusion: family environment is making an extremely significant impact on students' mathematics achievement
- More wealthy families are able to devote more resources for the education of their children which will contribute to their mathematics achievement
- And if a part of those resources is devoted to numeracy activities and tasks when the children were young, then it will somehow contribute to their mathematics achievement when they reached P4

Family income, parents' education level and jobs

$7,000 or below		$7,001 - $14,000		$14,001 - $21,000		$21,001 - $28,000		$28,001 - $36,000		$36,000 or above	
% of Students	Average Achievement	% of Students	Average Achievement	% of Students	Average Achievement	% of Students	Average Achievement	% of Students	Average Achievement	% of Students	Average Achievement
10	596 ✪	31	601 ✪	18	605 ✪	10	612 ✪	9	613 ✪	22	619

Finished University or Higher		**Finished Post-secondary Education**		**Finished Upper Secondary**		**Finished Lower Secondary Education or Less**	
% of Students	Average Achievement	% of Students	Average Achievement	% of Students	Average Achievement	%e of Students	Average Achievement
18	623	8	612 ✪	40	608 ✪	34	599 ✪

✪ Difference statistically significant with "Finished University or Higher"

Father's Job (Mathematics)

Never Worked		Small Business Owner		Clerk and Sales Worker		Skilled Worker		General Laborers		Professional and Associate Professional	
% of Students	Average Achievement	% of Students	Average Achievement	% of Students	Average Achievement	% of Students	Average Achievement	% of Students	Average Achievement	% of Students	Average Achievement
1	593 ✪	13	611 ✪	15	607 ✪	33	604 ✪	9	604 ✪	29	619

Mother's Job (Mathematics)

Never Worked		Small Business Owner		Clerk and Sales Worker		Skilled Worker		General Laborers		Professional and Associate Professional	
% of Students	Average Achievement	% of Students	Average Achievement	% of Students	Average Achievement	% of Students	Average Achievement	% of Students	Average Achievement	% of Students	Average Achievement
18	612 ✪	7	604 ✪	49	607 ✪	2	613 ✪	7	604 ✪	18	619

✪ Difference statistically significant with "Professional and Associate Professional"

D. Schools students attend

- The type of school students attend (finance type, gender type, school sponsoring body the school belongs to) make a lot of difference to the student's mathematics achievement.
- Primary 4 students attending private and Direct Subsidy Scheme (DSS) schools outperforming students attending other types of schools; students attending single-gender schools outperforming those in co-educational schools; and students in schools run by large and Christian School Sponsoring Bodies (SSBs) outperforming their counterparts
- For Secondary 2, students attending Government schools, single-gender schools, and schools operated by Christian SSBs perform better.

Grade 4

*SSB = School Sponsoring Body

Finance Type	Maths score
Government	600
Aided	599
DSS	620 *
Private	626 **
Average	602

Gender Type	Maths score
Boys	616 *
Girls	620 **
Co-educational	601
Average	602

SSB	Maths score
Mega Government	607* 600**
Large	611***
Medium	600 **
Small	596
Average	602

SSB*	Maths score
Catholic	609*
Protestant	609*
Oriental religions	545
No religion	596 **
Average	602

Grade 8

Financial Type	Maths score
Band 1	638 *
Band 2	554**
Band 3	523
Average	587

Gender Type	Maths score
Boys	618 *
Girls	605 **
Co-educational	575
Average	587

SSB	Maths score
Mega	590*
Government	610**
Large	578***
Medium	564
Small	592 *
Average	587

SSB	Maths score
Catholic	589*
Protestant	599**
Oriental religions	551
No religion	577 ***
Average	587

Some reflections on education equity

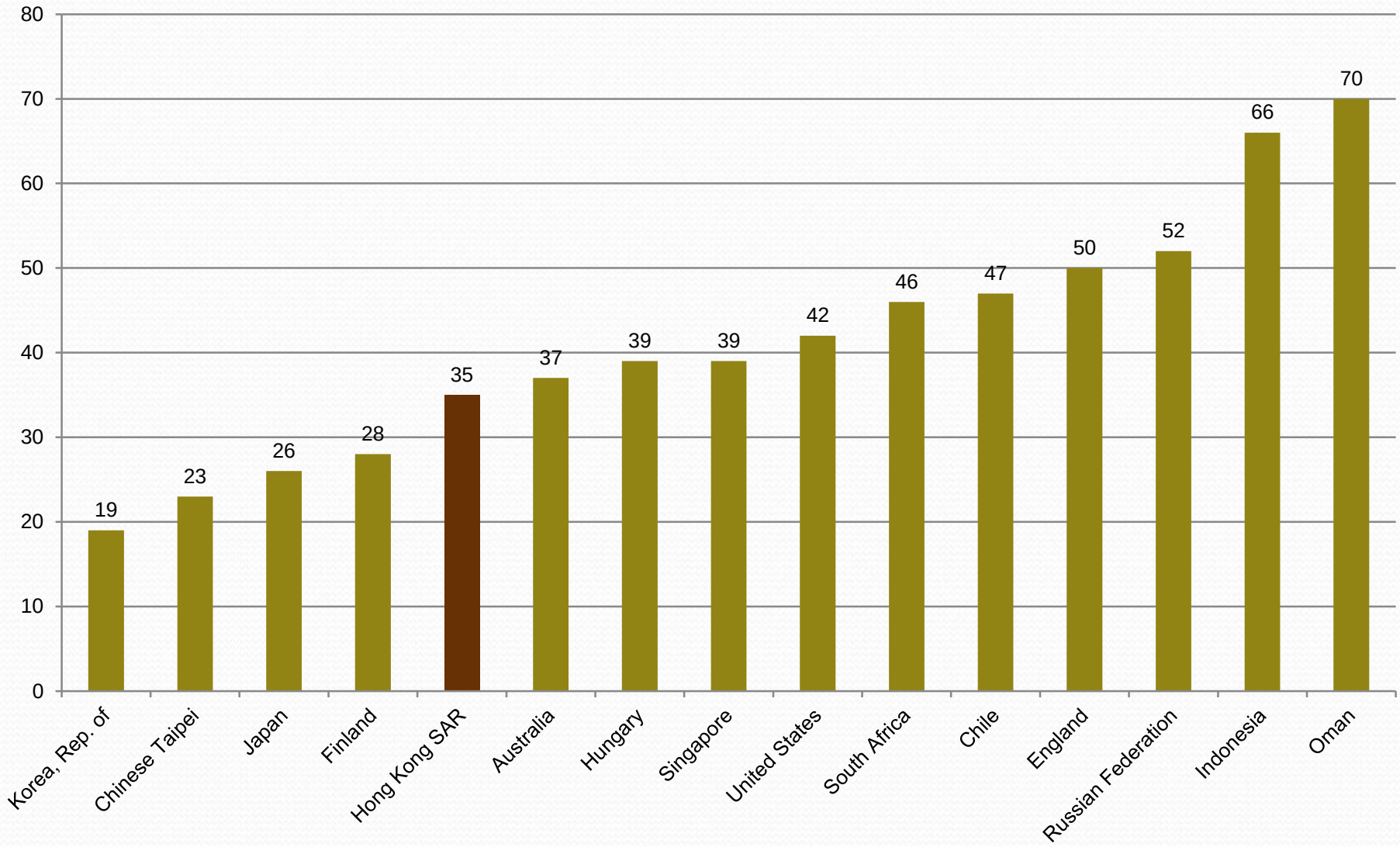
- How well has Hong Kong been addressing the issue of equity?
- The influence of SES on student achievement is a universal phenomenon – is Hong Kong better off or worse than other systems?
- Can anything be done to alleviate the influence of SES?
- School sponsoring body – a major characteristic of the Hong Kong system – what are the pros and cons?
- Students in government secondary schools outperformed their counterparts, but students in government primary schools did not do as well as their counterparts in private and DSS primary schools, and in schools belonging to mega and large SSBs, why?
- Students in Protestant schools, performed better than schools of other religious affiliations - what traditions these schools have established which have enabled their students to perform well in mathematics?
- Government/Subsidised /DSS schools – is this a fair school system

6. Attitudes of students towards mathematics and mathematics learning

- Students' attitudes are an important component of the attained curriculum, since in all school systems, students' positive attitudes are one of the goals of education
- In this era when life-long learning is so much stressed, some people think that a positive attitude is even more important than attaining high test scores
- A positive attitude will motivate students to continue to learn even after they have left school
- So we should care about students' attitude towards learning, not just their achievement

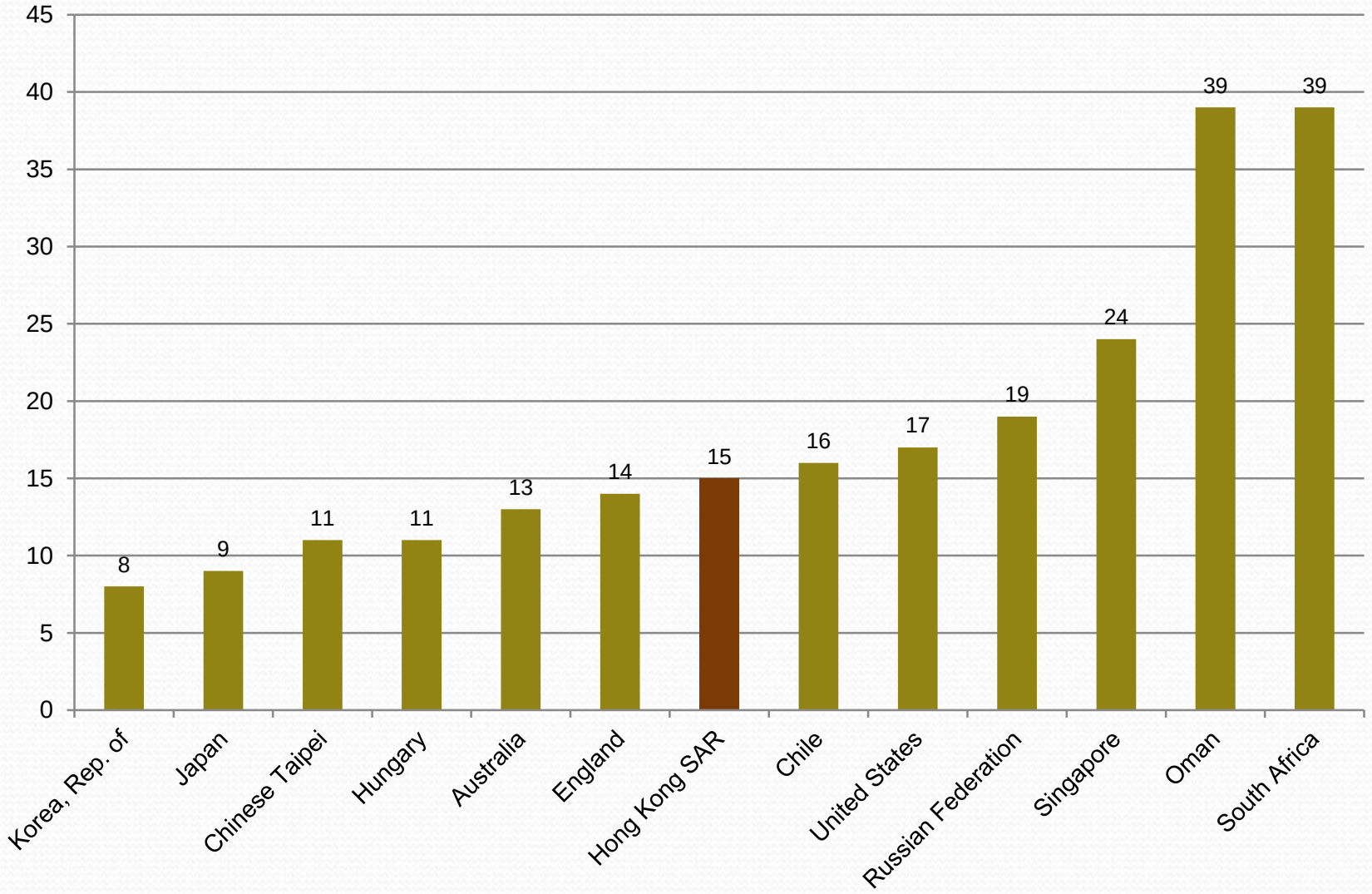
Grade 4: Students like learning mathematics

(international average = 46%)

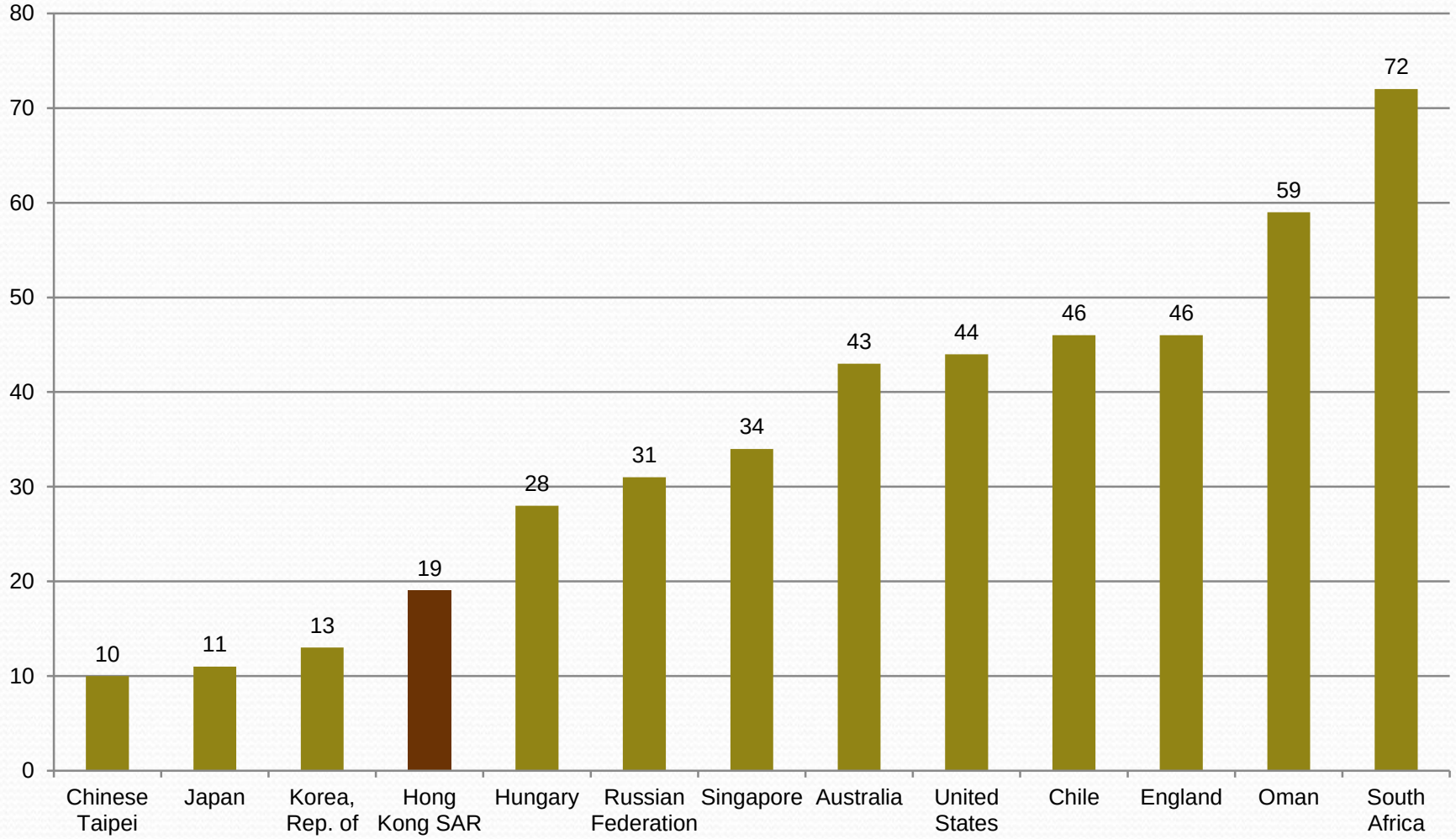


Grade 8: Students like learning mathematics

(international average = 22%)

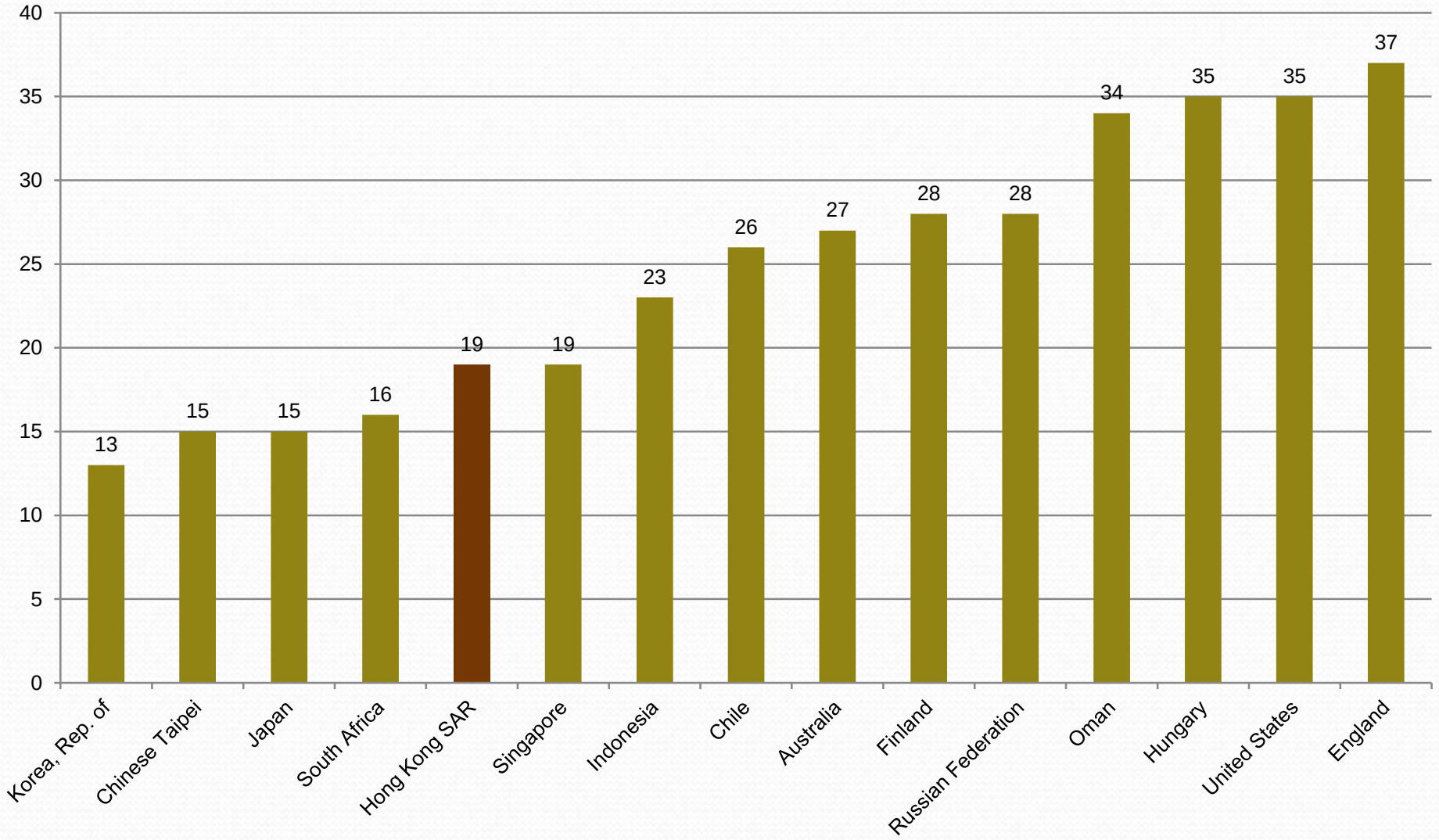


Grade 8: Students valuing mathematics (international average = 42%)



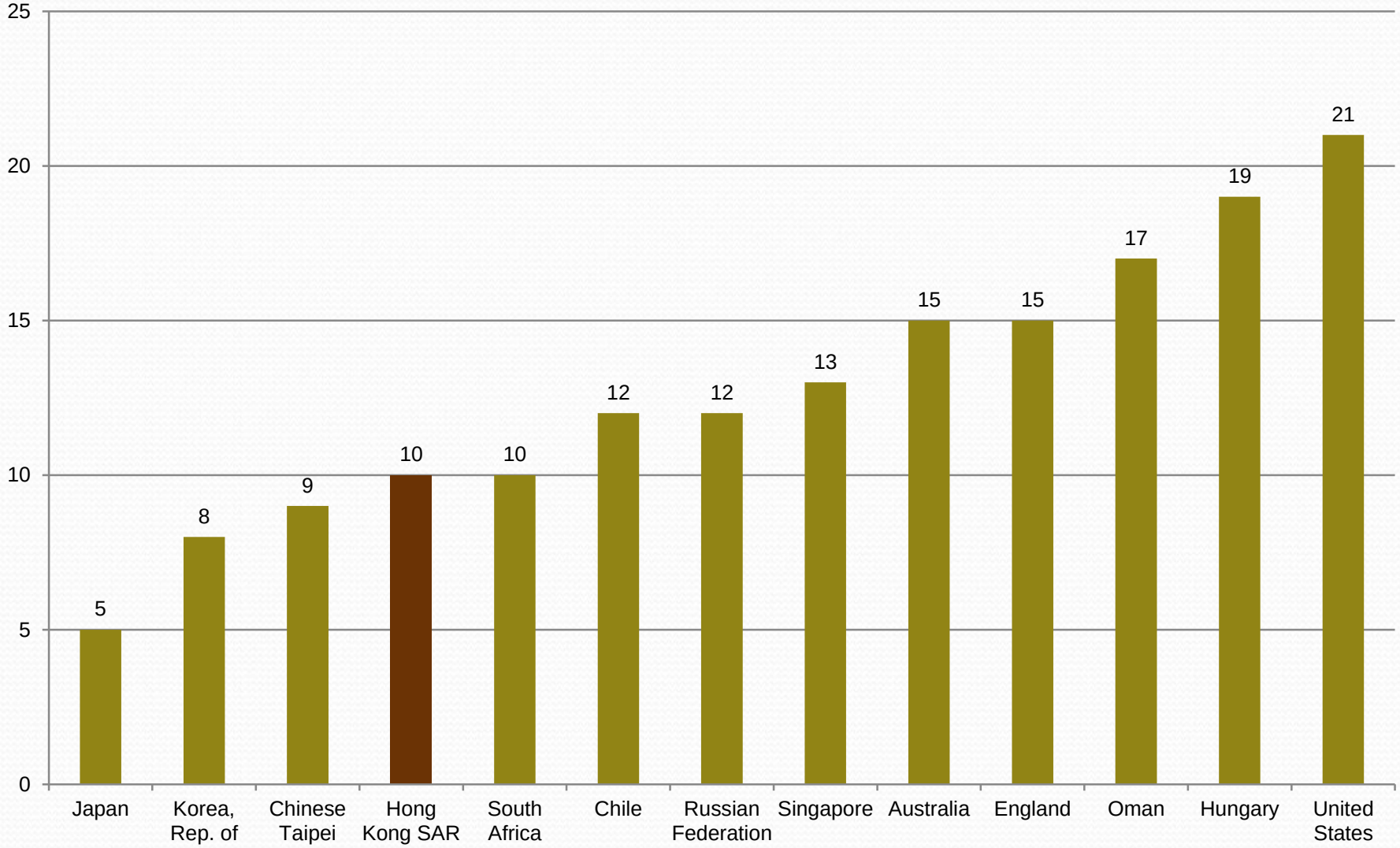
Grade 4: Students' confidence in mathematics

(international average = 32%)



Grade 8: Students' confidence in mathematics

(international average = 14%)



What price have we paid for high achievement?

- Students' physical health?
- Students' interest and development of hobbies?
- Students' enjoyment of school life?
- Students' enjoyment of family life?

Conclusion

- International studies are important for answering questions about the effectiveness of a school system that cannot be answered by research within a country
- Because of the nature and the limitations of these studies, we should be very careful in using results of these studies
- Suggesting drastic changes in education policies based on results of these studies without due consideration of the nature and limitations of these studies, as well as of the cultural differences among countries, may be misleading and even harmful
- Education is a complex endeavour – we cannot expect international studies to produce answers for all our national problems in education!

Acknowledgement

The Hong Kong component of TIMSS is funded by the Education Bureau of Hong Kong (EDB), but the views expressed in this presentation are those of the presenter and do not necessarily represent the views of EDB

Thank you very much for your attention!

My e-mail address:

frederickleung@hku.hk