Comparative Study of Mathematics Classrooms

- What can be learned from the TIMSS 1999 Video Study

Frederick K.S. Leung
The University of Hong Kong
APEC-Tsukuba Conference, Tokyo, Japan
16 January, 2006

Introduction

- Students from East Asian countries have consistently outperformed their counterparts in the West in comparative studies such as TIMSS and PISA
- The high achievements do not seem to be accompanied by correspondingly positive attitudes towards mathematics
- What accounts for the high achievement of East Asian students despite their negative attitudes towards mathematics?
- Since students learn most of their knowledge in the classroom, it is reasonable to expect that the instruction they receive should be a major factor in influencing their achievement

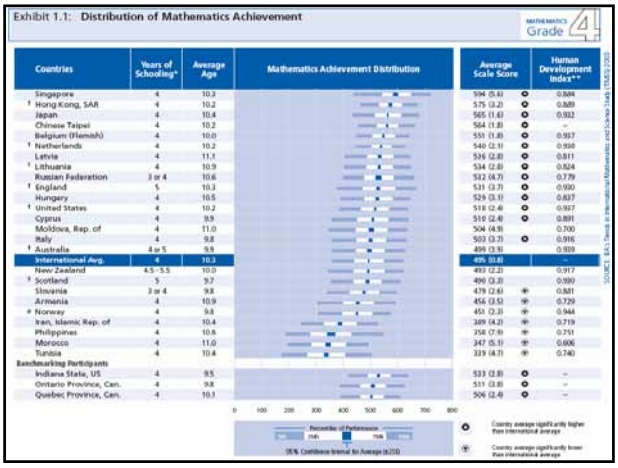


Exhibit 1.1: Distribution of Mathematics Achievement

MATHEMATICS Grade 4

Countries	Years of Schooling*	Average Age	Mathematics Achievement Distribution	Average Scale Score	Human Development Index**
Singapore	4	10.3		594 (5.6)	0.884
† Hong Kong, SAR	4	10.2		575 (3.2)	0.889
Japan	4	10.4		565 (1.6)	0.932
Chinese Taipei	4	10.2		564 (1.8)	–
Belgium (Flemish)	4	10.0		551 (1.8)	0.937
† Netherlands	4	10.2		540 (2.1)	0.938
Latvia	4	11.1		536 (2.8)	0.811
† Lithuania	4	10.9		534 (2.8)	0.824
Russian Federation	3 or 4	10.6		532 (4.7)	0.779
† England	5	10.3		531 (3.7)	0.930
Hungary	4	10.5		529 (3.1)	0.837
† United States	4	10.2		518 (2.4)	0.937
Cyprus	4	9.9		510 (2.4)	0.891
Moldova, Rep. of	4	11.0		504 (4.9)	0.700
Italy	4	9.8		503 (3.7)	0.916
† Australia	4 or 5	9.9		499 (3.9)	0.939
International Avg.	4	10.3		495 (0.8)	–
New Zealand	4.5 - 5.5	10.0		493 (2.2)	0.917
‡ Scotland	5	9.7		490 (3.3)	0.930
Slovenia	3 or 4	9.8		479 (2.6)	0.881
Armenia	4	10.9		456 (2.5)	0.729
# Norway	4	9.8		453 (2.3)	0.944
Iran, Islamic Rep. of	4	10.4		389 (4.2)	0.719
Philippines	4	10.6		358 (7.9)	0.751
Morocco	4	11.0		347 (5.1)	0.606
Tunisia	4	10.4		339 (4.7)	0.740
Benchmarking Participants					
Indiana State, US	4	9.5		533 (2.8)	–
Ontario Province, Can.	4	9.8		511 (2.8)	–
Quebec Province, Can.	4	10.1		506 (2.4)	–

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

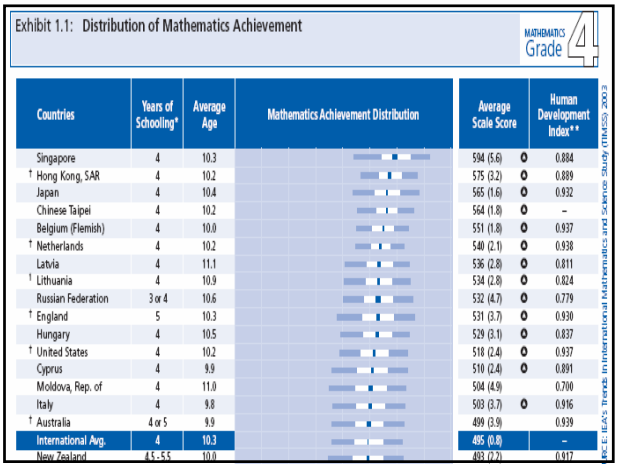


Exhibit 1.1: Distribution of Mathematics Achievement

MATHEMATICS Grade 4

Countries	Years of Schooling*	Average Age	Mathematics Achievement Distribution	Average Scale Score	Human Development Index**
Singapore	4	10.3		594 (5.6)	0.884
† Hong Kong, SAR	4	10.2		575 (3.2)	0.889
Japan	4	10.4		565 (1.6)	0.932
Chinese Taipei	4	10.2		564 (1.8)	–
Belgium (Flemish)	4	10.0		551 (1.8)	0.937
† Netherlands	4	10.2		540 (2.1)	0.938
Latvia	4	11.1		536 (2.8)	0.811
† Lithuania	4	10.9		534 (2.8)	0.824
Russian Federation	3 or 4	10.6		532 (4.7)	0.779
† England	5	10.3		531 (3.7)	0.930
Hungary	4	10.5		529 (3.1)	0.837
† United States	4	10.2		518 (2.4)	0.937
Cyprus	4	9.9		510 (2.4)	0.891
Moldova, Rep. of	4	11.0		504 (4.9)	0.700
Italy	4	9.8		503 (3.7)	0.916
† Australia	4 or 5	9.9		499 (3.9)	0.939
International Avg.	4	10.3		495 (0.8)	–
New Zealand	4.5 - 5.5	10.0		493 (2.2)	0.917

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

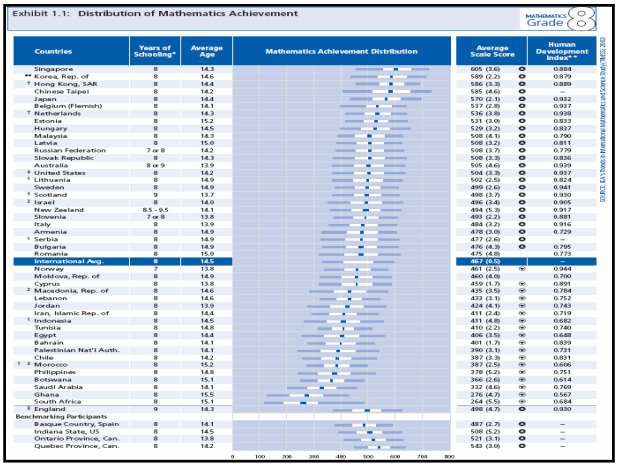
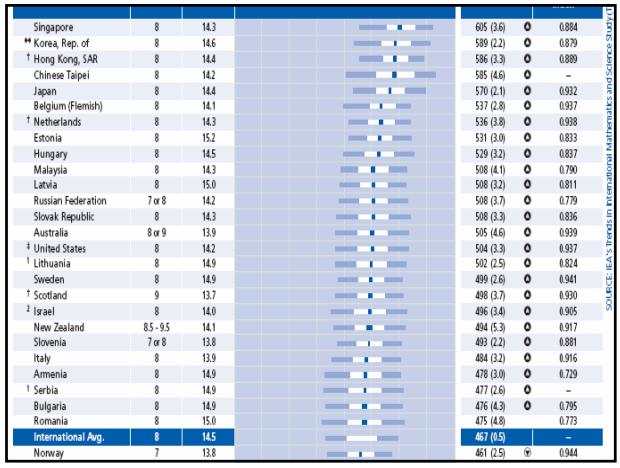


Exhibit 1.1: Distribution of Mathematics Achievement

MATHEMATICS Grade 8

Countries	Years of Schooling*	Average Age	Mathematics Achievement Distribution	Average Scale Score	Human Development Index**
Singapore	8	14.3		605 (3.6)	0.884
** Korea, Rep. of	8	14.6		589 (2.2)	0.879
† Hong Kong, SAR	8	14.4		586 (3.3)	0.889
Chinese Taipei	8	14.2		585 (4.6)	–
Japan	8	14.4		570 (2.1)	0.932
Belgium (Flemish)	8	14.1		537 (2.8)	0.937
† Netherlands	8	14.3		536 (3.8)	0.938
Estonia	8	15.2		531 (3.0)	0.833
Hungary	8	14.5		529 (3.2)	0.837
Malaysia	8	14.3		508 (4.1)	0.790
Latvia	8	15.0		508 (3.2)	0.811
Russian Federation	7 or 8	14.2		508 (3.7)	0.779
Slovak Republic	8	14.3		508 (3.3)	0.836
Australia	8 or 9	13.9		505 (4.6)	0.939
‡ United States	8	14.2		504 (3.3)	0.937
† Lithuania	8	14.9		502 (2.5)	0.824
Sweden	8	14.9		499 (2.6)	0.941
† Scotland	9	13.7		498 (3.7)	0.930
² Israel	8	14.0		496 (3.4)	0.905
New Zealand	8.5 - 9.5	14.1		494 (5.3)	0.917
Slovenia	7 or 8	13.8		493 (2.2)	0.881
Italy	8	13.9		484 (3.2)	0.916
Armenia	8	14.9		478 (3.0)	0.729
† Serbia	8	14.9		477 (2.6)	–
Bulgaria	8	14.9		476 (4.3)	0.795
Romania	8	15.0		475 (4.8)	0.773
International Avg.	8	14.5		467 (0.5)	–
Norway	7	13.8		461 (2.5)	0.944
Moldova, Rep. of	8	14.9		460 (4.0)	0.700
Cyprus	8	13.8		459 (1.7)	0.891
² Macedonia, Rep. of	8	14.6		435 (3.5)	0.784
Lebanon	8	14.6		433 (3.1)	0.752
Jordan	8	13.9		424 (4.1)	0.743
Iran, Islamic Rep. of	8	14.4		411 (2.4)	0.719
† Indonesia	8	14.5		411 (4.8)	0.682
Tunisia	8	14.8		410 (2.2)	0.740
Egypt	8	14.4		406 (3.5)	0.688
Bahrain	8	14.1		401 (1.7)	0.839
Palestinian Nat'l Auth.	8	14.1		390 (3.1)	0.731
Chile	8	14.2		387 (3.3)	0.831
† ‡ Morocco	8	15.2		387 (2.5)	0.606
Philippines	8	14.8		378 (5.2)	0.751
Botswana	8	15.1		366 (2.6)	0.614
Saudi Arabia	8	14.1		332 (4.6)	0.769
Ghana	8	15.5		276 (4.7)	0.567
South Africa	8	15.1		264 (5.5)	0.684
‡ England	9	14.3		498 (4.7)	0.930
Benchmarking Participants					
Basque Country, Spain	8	14.1		487 (2.7)	–
Indiana State, US	8	14.5		508 (5.2)	–
Ontario Province, Can.	8	13.8		521 (3.1)	–
Quebec Province, Can.	8	14.2		543 (3.0)	–

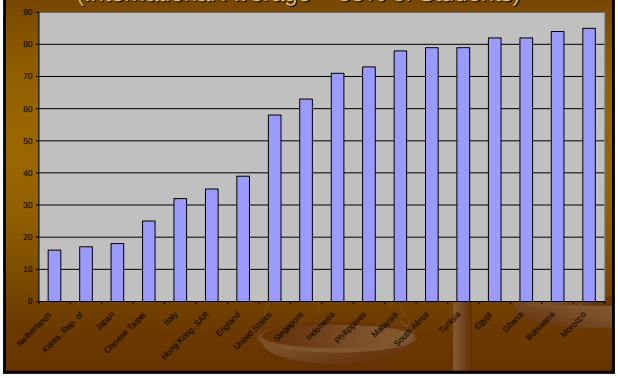
SOURCE: IEA's Trends in International Mathematics and Science Study (TIMSS) 2003


Singapore	8	14.3		605 (3.6)	0.884
** Korea, Rep. of	8	14.6		589 (2.2)	0.879
† Hong Kong, SAR	8	14.4		586 (3.3)	0.889
Chinese Taipei	8	14.2		585 (4.6)	–
Japan	8	14.4		570 (2.1)	0.932
Belgium (Flemish)	8	14.1		537 (2.8)	0.937
† Netherlands	8	14.3		536 (3.8)	0.938
Estonia	8	15.2		531 (3.0)	0.833
Hungary	8	14.5		529 (3.2)	0.837
Malaysia	8	14.3		508 (4.1)	0.790
Latvia	8	15.0		508 (3.2)	0.811
Russian Federation	7 or 8	14.2		508 (3.7)	0.779
Slovak Republic	8	14.3		508 (3.3)	0.836
Australia	8 or 9	13.9		505 (4.6)	0.939
‡ United States	8	14.2		504 (3.3)	0.937
† Lithuania	8	14.9		502 (2.5)	0.824
Sweden	8	14.9		499 (2.6)	0.941
† Scotland	9	13.7		498 (3.7)	0.930
² Israel	8	14.0		496 (3.4)	0.905
New Zealand	8.5 - 9.5	14.1		494 (5.3)	0.917
Slovenia	7 or 8	13.8		493 (2.2)	0.881
Italy	8	13.9		484 (3.2)	0.916
Armenia	8	14.9		478 (3.0)	0.729
† Serbia	8	14.9		477 (2.6)	–
Bulgaria	8	14.9		476 (4.3)	0.795
Romania	8	15.0		475 (4.8)	0.773
International Avg.	8	14.5		467 (0.5)	–
Norway	7	13.8		461 (2.5)	0.944

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

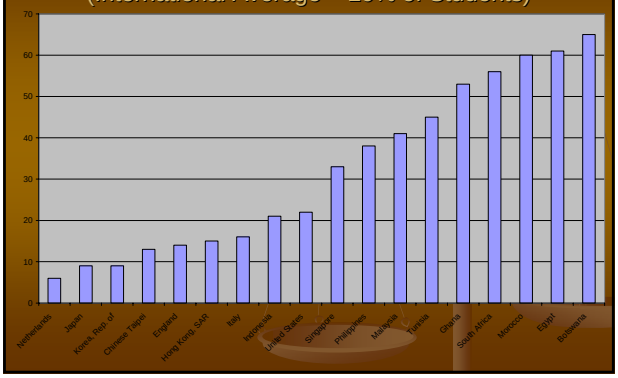
Students' Valuing of Mathematics (TIMSS 2003)

(International Average = 55% of Students)



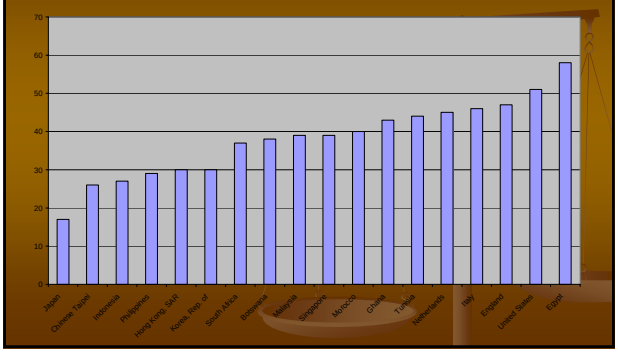
Students' Enjoyment of Mathematics (TIMSS 2003)

(International Average = 29% of Students)



Students' Self-Confidence in Learning Mathematics (TIMSS 2003)

(International Average = 40% of Students)



This presentation

- TIMSS 1999 Video Study data for Hong Kong analyzed to seek explanation of students' high achievement in mathematics
- Methodological issues on comparative classroom studies are discussed
- Results of the Study interpreted with reference to the East Asian culture
- Implications of the findings of the study drawn for mathematics education in East Asian and other countries

TIMSS 1999 Video Study (Math)

Goal:

Examine instructional practices in eighth-grade mathematics across seven countries: Australia, Czech Republic, Hong Kong SAR, Japan*, Netherlands, Switzerland, United States

* The 1995 Japanese data were re-analyzed using the 1999 methodology in some of the analysis

Sampling

- National probability sample of a target of 100 schools each country
- A **Video Survey**
- One lesson per teacher per school
- Sampled across the school year
- Standardized camera procedures
- 638 lessons (e.g., Japan: 50; the Netherlands: 78; Hong Kong: 100; Switzerland: 140)

Data Coding and Analysis

- Standardized camera procedures
- International team developed codes, apply to video data
- Fluently bilingual coders applied 45 codes in 7 coding passes to each videotaped lesson
- 3 marks (in-point, out-point, and category) evaluated and included in the measures of reliability
- If, after numerous attempts, reliability measures fell below the minimum acceptable standard, the code was dropped from the study
- Danger: fine grained analysis breaks lessons down into minute constituent parts, but the parts may not fit with each other to form back a meaningful picture of the lesson

The Mathematics Quality Analysis Group

- A specialist group in mathematics and teaching mathematics at the post-secondary level reviewed a randomly selected subset of 120 lessons (20 lessons from each country except Japan)
- The international video coding team created expanded lesson tables for each lesson in this subset.
- The tables included details about the classroom interaction, the nature of the math problems worked on, mathematical generalizations, and other relevant information
- The tables were "country-blind," with all indicators that might reveal the country removed

A. Instructional Practices in Hong Kong as Portrayed by the Analysis of the Codes

Whole-class interaction dominated

"Comparing across countries, eighth-grade mathematics lessons in Hong Kong SAR spent a greater percentage of lesson time in public interaction (75 percent) than those in the other countries, except the United States." (pp. 54-55)

Average percentage of lesson time devoted to public and private interaction

Country	Public interaction	Private interaction	Others
Australia	52	48	0
Czech Republic	61	21	18
Hong Kong	75	20	5
Japan	63	34	3
Netherlands	44	55	1
Switzerland	54	44	1
United States	67	32	1

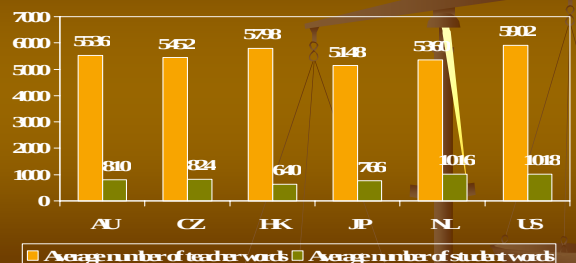
Teacher talked most of the time

"Hong Kong SAR eighth-grade mathematics teachers spoke significantly more words relative to their students (16:1) than did teachers in Australia (9:1), the Czech Republic (9:1), and the United States (8:1)." (p. 109, Chapter 5)

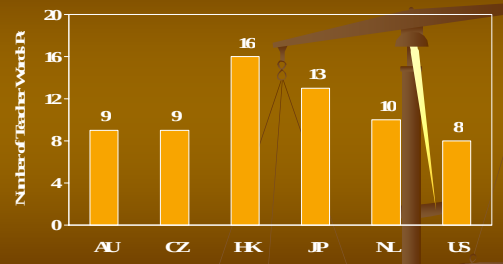
- Factor in the large class size (37) of Hong Kong, students' reticence even more striking
- Ratio of 16 to 1 is in effect a ratio of nearly 600:1 as far as an individual student in concerned

(Average class size ranged from 19 (Switzerland) to 27 (Australia)).

Average number of teacher and student words per lesson



Average number of teacher words to every one student word per lesson

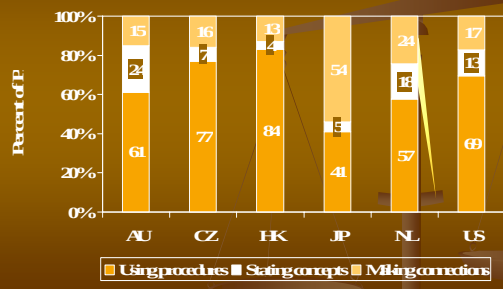


Students solve procedural problems unrelated to real-life following prescribed methods

Nature of Problem Statements

"Hong Kong SAR lessons contained a larger percentage of problem statements classified as using procedures (84 percent) than all the other countries except the Czech Republic (77 percent)." (p. 98)

Average percentage of problems per lesson of each problem statement type



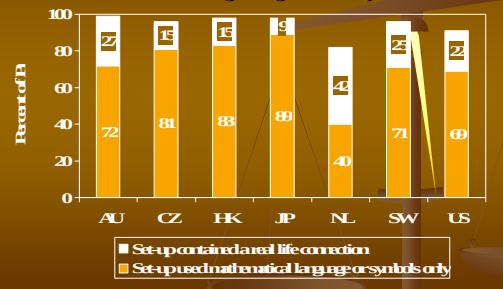
Contexts of the problems

- Many mathematics educators argue that mathematics problems presented within real-life contexts make mathematics more meaningful and hence more interesting for students

Choice of solution methods

- Mathematics educators usually think that to enhance students' problem solving ability, they should be encouraged to solve the same problem with different methods

Average percentage of problems per lesson set up with a real life connection or with mathematical language or symbols only



Average percentage of problems per lesson and percentage of lessons with at least one problem in which students had a choice of solution methods

Country	Average percent of problems with a choice of solution methods	Percent of lessons with at least one problem with a choice of solution methods
Australia	8	25
Czech Republic	4	20
Hong Kong	3	17
Japan	15	31
Switzerland	7	24
United States	9	45

Instructional practices as portrayed by the analysis of the codes

- Whole-class interaction dominated
- Teacher talked most of the time
- Problems solved by students:
  - procedural problems
  - unrelated to real-life
  - following prescribed methods

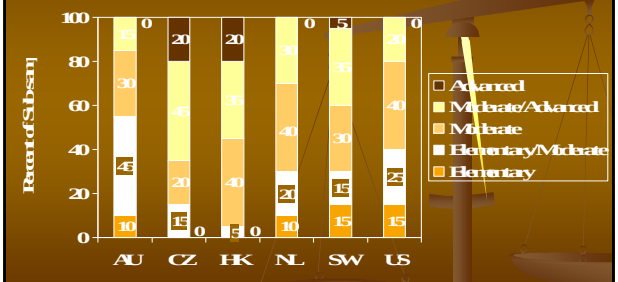
Not conducive to quality teaching and learning!

B. Quality of Content as judged by the Math Quality Analysis Group

More advanced content

"the ratings for countries with the most advanced (5) to the most elementary (1) content in the sub-sample of lessons, were the Czech Republic and Hong Kong SAR (3.7), Switzerland (3.0), the Netherlands (2.9), the United States (2.7), and Australia (2.5)" (p. 191)

Percentage of lessons in sub-sample at each content level



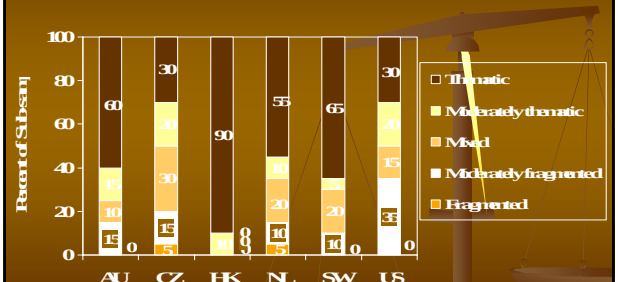
Lesson more coherent

Coherence = "the (implicit and explicit) interrelation of all mathematical components of the lesson" (p.196)

90% of the Hong Kong lessons were judged to be thematically coherent, with the remaining 10% moderately thematically coherent

Compares favorably with Czech Republic and the United States, where only 30% of the lessons were judged to be thematically coherent

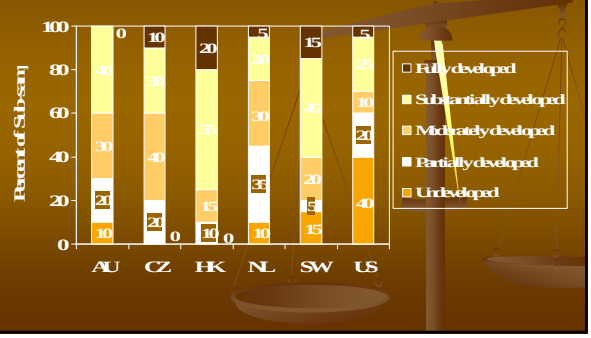
Percentage of lessons in sub-sample rated at each level of coherence



More fully developed presentation

- Presentation = "the extent to which the lesson included some development of the mathematical concepts or procedures".
- Development required that mathematical reasons or justifications were given for the mathematical results presented or used.
- Presentation ratings took into account the quality of mathematical arguments.
- Higher ratings meant that sound mathematical reasons were provided by the teacher (or students) for concepts and procedures.
- Mathematical errors made by the teacher reduced the ratings

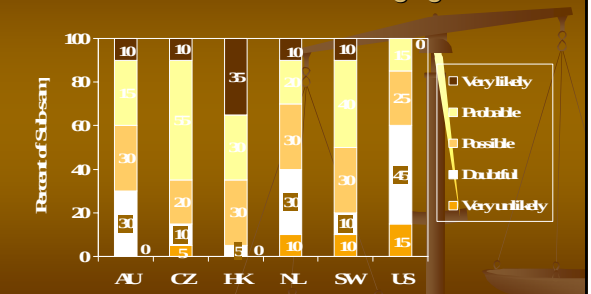
Percentage of lessons in sub-sample rated at each level of presentation



Students more likely to be engaged

- Student engagement = "the likelihood that students would be actively engaged in meaningful mathematics during the lesson".
- A rating of very unlikely (1) indicated a lesson in which students were asked to work on few of the problems and those problems did not appear to stimulate reflection on math concepts or procedures.
- A rating of very likely (5) indicated a lesson in which students were expected to work actively on, and make progress solving, problems that appeared to raise interesting mathematical questions for them and then to discuss their solutions with the class.

Percentage of lessons in sub-sample rated at each level of student engagement



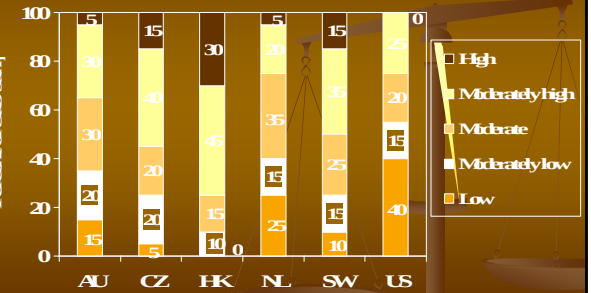
Overall quality

Overall quality judgment:

"the opportunities that the lesson provided for students to construct important mathematical understandings" (p. 199)

"the relative standing of Hong Kong SAR was consistently high …." (p. 200)

Percentage of lessons in sub-sample rated at each level of overall quality



General ratings for each overall dimension of content quality of lessons

Quality of the Content as judged by the Math Quality Analysis Group

- Relatively advanced content
- More deductive reasoning
- More coherent
- More fully developed presentation
- Students are more engaged, and
- Overall quality is high

A more positive picture!

Two different pictures?

- How do we reconcile the apparent inconsistency between the instructional practices as reflected by the two different analyses of the same data set?
- In the first picture, instructional practices were portrayed through objectively coding and summarizing the activities that happened in the classroom
- In the second picture, the quality of content was judged by the Mathematics Quality Analysis Group based on their expertise and experience

Qualitative analysis

- Small sample size (20 for each country)
- The sub-sample "might not be representative of the entire sample or of eighth-grade mathematics lessons in each country" (Hiebert *et al*, 2003: 190)
- Results not very reliable
- Analysis relied on the judgement of a group of experts: "rater-dependent"
- Another group of experts may arrive at rather different conclusions

Quantitative analysis

- Coding (e.g. number of words spoken by the teachers and students) low-inferenced, hence relatively objective
- Provide relevant information the interactions in the lessons
- But quality of what the teachers and students say far more important that how much they say
- To determine what they said is significant or not requires judgement based on experience, hence subjective
- Results of quantitative analysis may be highly reliable but not necessarily very meaningful

Reliability and validity of video data

- Inherent trade-off between reliability and validity?
- To get reliable data, we restrain from making inferences, and hence lose out in validity
- To increase validity, we make subjective judegment, resulting in low reliability
- Rather like Heisenberg's Principle of Uncertainty in physics
- To get reliable and objective information, we lose out in the meaningfulness or validity
- But qualitative analysis involves judgement of "experts", and different experts may yield different results. Hence results unreliable

Which is the "real" picture in Hong Kong?

- Expert judgment unreliable?
- Quantitative analysis fail to reveal the subtlety of the complexity of classroom teaching?
- Do we prefer a reliable description of the activities in the classroom, or tolerate lack of reliability and learn from experts' view on the quality of teaching and learning?
- The crux of the matter: in determining the quality of teaching, should we rely on objective summary of data, or should we rely on the subjective judgment of the experts?
- Synthesis of the two gives nearer to the reality

Pragmatic philosophy in the East Asian Culture

- Replication of Ma's study (Ma, 1999) in Hong Kong and Korea (Leung and Park, 2002):
  - teachers in the study were in general competent in mathematics
  - they deliberately taught in a procedural manner for pedagogical reasons and for the sake of efficiency
  - teachers believe that it would be inefficient or even confusing for school children to be exposed to rich concepts instead of clear and simple procedures

Virtue of humility or modesty

Children from these countries are taught from when they are young that one should not be boastful. …. one's confidence and self image are something that is reinforced by one's learned values, and if students are constantly taught to rate themselves low, they may internalize the idea to result in really low confidence. Furthermore, the competitive examinations systems coupled with the high expectations for student achievement in these countries have left a large number of students classified as failures in their system, and these repeated experiences of a sense of failure may have further reinforced this lack of confidence.

(Leung, 2002: 106)

Why East Asian students perform well?

- Negative correlation between students' confidence in mathematics and their achievement is expected
- Stress on diligence and practice
- Attributes success more to effort than to innate ability
- Confucian values: emphasize strongly on the importance of education and a high expectation for students to achieve
- Learning or studying is considered a serious endeavour, and students are expected to put in hard work and perseverance
- Reinforced by a long and strong tradition of publication examination, which acts as a further source of motivation for learning

Implications for East Asian countries

- Results of the quantitative analysis may prompt a call for radical changes in the instructional practice in Hong Kong classrooms and East Asia
- Results from the qualitative analysis present a different picture
- Qualitative results more consistent with results of the achievement data
- Quantitative analysis done using more objective methodology utilizing a more representative dataset
- The two sets of results complement each other in giving a picture closer to the reality of the East Asian classroom

- Strength of East Asian instructional practices
  - teaching mathematics in a coherent manner
  - Fully develop the lessons and keep students engaged
- Expectation that students should learn a relatively advanced level of content with an appropriate degree of emphasis on deductive reasoning
- Reducing content difficulty to make mathematics more accessible to the general student population an endless retreat
- Teachers should make the lessons lively and enjoyable to induce students' interest in mathematics rather than in the activities

Weaknesses

- Dominance of teacher talk: not the best kind of activities for effective mathematics learning
- The fact that the majority of the problems student solve are unrelated to real-life not consistent with the ideal of a good mathematics education
- Challenge for mathematics educators in East Asia: promote student participation in meaningful learning without compromising their strengths in instructional practices

Implications for other countries

- Students' achievement in TIMSS should be viewed in conjunction with their attitudes towards mathematics and mathematics learning
- Enjoyment of study part of the aims of education in all cultures
- Negative attitudes part of the attained curriculum, and educators should be alarmed
- High student achievement should not relegate efforts to promote students' interest
- Simple transplant of educational policies from high achieving countries to low achieving ones would not work

- One cannot transplant the practice without regard to the cultural differences
- Culture by definition evolves slowly and stably with the passage of long periods of time, and there is simply no quick transformation of culture
- Need to identify not only the superficial differences in educational practice, but the intricate relationship between educational practice and the underlying culture
- Through studying these relationships, we understand the interaction between educational practices and culture
- Through identifying the commonality and differences of both the educational practices and the underlying cultures, we may then determine how much can or cannot be borrowed from another culture

Conclusion

- Public interest of international comparative studies usually focused on the relative position of countries in the league tables generated
- Provides impetus for curriculum changes, but very often, changes are made without careful consideration of the context within which the achievement and classroom instructions are situated
- The purpose of international studies is not competition
- Results of TIMSS 1999 Video Study should not be used to justify classroom practices of some countries
- The rich dataset generated serves as mirrors for us to better understand our system
- Policy changes should take culture into account

Thank you very much for your attention!

My e-mail address:
frederickleung@hku.hk