COMPARATIVE STUDY OF MATHEMATICS CLASSROOMS

– WHAT CAN BE LEARNED FROM THE TIMSS 1999 VIDEO STUDY[1]

Frederick K.S. Leung

The University of Hong Kong

The Third International Mathematics and Science Study (TIMSS) 1999 Video Study aims at describing and comparing eighth-grade mathematics teaching practices among seven countries in order to identify similar or different classroom features. Since East Asian students have consistently performed well in recent international studies of mathematics achievement, this paper intends to analyze the TIMSS Video Study data for the East Asian country of Hong Kong in order to see whether there are classroom practices that can be used to explain students' high achievement in mathematics. The data analysis however yields conflicting results. While a qualitative analysis of the data shows that the quality of mathematics teaching in Hong Kong is high, a quantitative analysis of the same data shows that teaching in Hong Kong is rather traditional and teacher-centred. The conflicting results point to the complexity in interpreting video data on classroom practices and of achievement data in international studies. The results are then interpreted with respect to the underlying cultural values in East Asia, and implications for methodology in analyzing video data, as well as for educational reform in East Asian countries and other countries are discussed.

Introduction

Students from East Asian countries [2] have consistently outperformed their counterparts in the West in international comparative studies of mathematics achievement such as the *Third International Mathematics and Science Study* (TIMSS[3]) (Beaton *et al,* 1996; Mullis *et al,* 1997; Mullis *et al,* 2000; Mullis *et al,* 2004) and the OECD *Program for International Student Assessment* (PISA) (OECD, 2001; 2003; 2004). However, the high achievements of East Asian students do not seem to have been accompanied by correspondingly positive attitudes towards mathematics (Leung, 2002). An obvious question to ask of such

---

[1] Paper to be delivered at the APEC-Tsukuba Conference, Tokyo, Japan, 16 January, 2006.

[2] East Asian "countries" in this paper refer to Chinese Taipei, Hong Kong, Japan, Korea and Singapore. Although some of them (e.g. Hong Kong) are not countries, for convenience the generic term "countries" will be used to refer to all participants in these international studies.

[3] TIMSS was renamed Trends in *International Mathematics and Science Study* starting from the 2003 Study.

international studies is what accounts for high achievement, and in particular, 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.

In this paper, the TIMSS 1999 Video Study data for the East Asian country of Hong Kong are analyzed to see whether there are classroom practices that can be used to explain students' high achievement in mathematics. Methodological issues related to comparative classroom studies are then discussed, and results of the Study are interpreted with reference to the East Asian culture. Finally some implications of the findings of the study are drawn for mathematics curriculum development in East Asian and other countries.

The TIMSS 1999 Video Study

The TIMSS 1999 Video Study (hereafter referred to as the Study) examined instructional practices in eighth grade mathematics for seven countries: Australia, Czech Republic, Hong Kong SAR, Japan, Netherlands, Switzerland, United States. The goals of the Study were to:

• describe and compare eighth-grade mathematics teaching across seven countries

• discover alternative ways to teach mathematics

• examine teaching in one's own country with fresh eyes, and

• create digital library of public use videos for teacher professional development

(Hiebert *et al*, 2003)

Japan did not collect video data for mathematics in 1999, but the Japanese data for the TIMSS 1995 video study were re-analyzed using the 1999 methodology in some of the analyses. For this reason, only the Hong Kong data will be highlighted for discussion below, since it is the only East Asian country for which data was collected in the 1999 Study.

Sampling, Data Collection and Analysis

<u>Sampling</u>

To obtain a representative sample of eighth-grade mathematics classrooms in each of the participating countries, a national probability sample[4] of a target of 100 schools was drawn in the Study. One mathematics class was then randomly selected from each of the schools, and only one lesson was videotaped for the sampled class. Including the 50 Japanese lessons videotaped in 1995, altogether 638 lessons were videotaped, ranging from 78 lessons (the Netherlands) to 140 lessons (Switzerland[5]) per country.

Since the eighth grade mathematics curriculum in the seven countries differs from each other, it has not been possible to match the content of the videotaped lessons in different countries. Instead, lessons were randomly selected across the school year so that they covered the content taught in the whole of the eighth grade in the country.

Data Coding and Analysis

Videotaping in all countries followed standardized camera procedures. Two cameras were used, with one camera focusing on the teacher and her interaction with students, and the other camera focusing on the whole class. All data from the seven countries were assembled together and analyzed by an international video coding team, advised by an expert group with members (known as national research coordinators) from each of the participating countries. Although the working language of the project was English, data analysis for individual countries was performed in the language used in the classrooms. Members of the international video coding team were all fluently bilingual (in the language used in the classrooms concerned and English) researchers, and working together they developed codes to apply to the video data. Three marks (i.e., the in-point, out-point, and category) for the codes were evaluated and included in the measures of reliability. For any code, if the reliability measures fell below the minimum acceptable standard after numerous attempts, it would then be dropped from the study. Altogether, 45 codes survived and were applied in seven coding passes to each of the videotaped lessons.

The Mathematics Quality Analysis Group

The quantitative analysis described above is fine grained, and allows details of the lessons to be captured. On the other hand, there is a danger that fine grained analysis would break down the lessons into minute constituent parts but the parts

---

[4] The sample was a Probability Proportional to Size (PPS) one, i.e., the probability that a school being chosen is proportional to the size of the school as measured by the number of eighth grade students in the school.

[5] In Switzerland, since there were three major languages of instruction, more schools were selected so that instructions across different language groups may be compared.

may not fit with each other to form back a meaningful picture of the lesson. For this reason, in addition to the quantitative analysis described above, a number of more qualitative analyses were performed. One such analysis was performed by an expert panel, known as the Mathematics Quality Analysis Group, comprising mathematicians and mathematics educators at the post-secondary level. The group reviewed a randomly selected subset of 120 lessons (20 lessons from each country except Japan[6]) and evaluated the quality of the lessons based on expanded "lesson tables" prepared by the international video coding team. The "lesson tables" contained detailed written descriptions of the lessons, including the classroom interaction, the nature of the mathematical problems worked on, goal statements, lesson summaries, and other relevant information. These lesson descriptions were examined "country-blind", with all indicators that might reveal the country removed.

Mathematics Classrooms in Hong Kong

A.   Instructional Practices as Portrayed by the Analysis of the Codes

Whole-class interaction dominated

In describing the kinds of teacher and students interaction in the seven countries, the Study defined five types of classroom interaction: public interaction, private interaction, student presents information, teacher presents information, and mixed private and public work. An analysis of the different types of interaction showed that the Hong Kong classroom was dominated by public or whole-class interaction. Three quarters of the lesson time was spent in public interaction while 20% of the lesson time was spent in private interaction (see Table 1 below). These represent the largest proportion of lesson time in public interaction and the smallest proportion of lesson time in private interaction among the seven countries. As the Study Report commented, "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." (Hiebert *et al*, 2003: 54-55).

---

[6] Since this same group of experts performed a similar analysis on the 1995 TIMSS Video data, which included the Japanese data, the 1999 Japanese data was not included in this analysis.

Table 1: Average Percentage of Lesson Time Devoted to Public and Private Interactions

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

What were the Hong Kong teachers and students doing during the whole-class interaction time? The Study recorded and calculated the number of words spoken by the teachers and the students in the lessons[7] as indication of the kind of interaction that took place. As can be seen from Figure 1 below, Hong Kong teachers spoke an average of about 5800 words per lesson while their students spoke only an average of 640 words. Compared to other countries in the study, Hong Kong teachers (together with the US teachers who spoke an average of about 5900 words per lesson) were the most talkative among the teachers in the participating countries. In contrast, Hong Kong students were the least talkative among the students in all the seven countries.
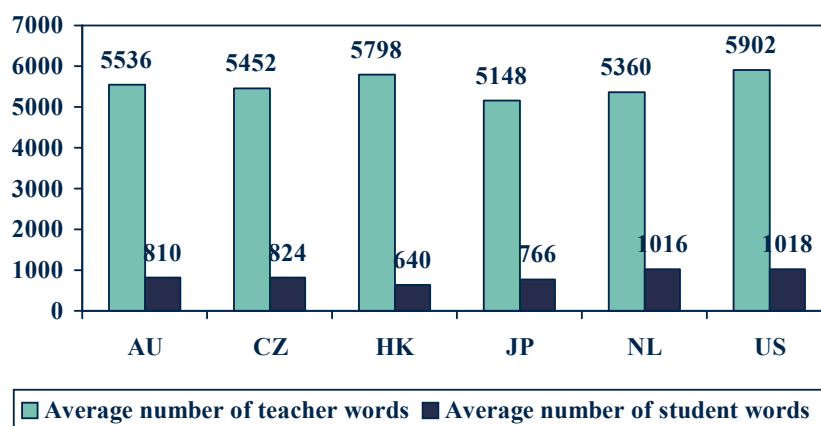

Figure 1: Average Number of Teacher and Student Words Per Lesson

Combining the two sets of figures in Figure 1, Hong Kong classrooms have the highest ratio of average number of words spoken by the teacher to those spoken by

---

[7] Since lesson duration varies across countries, the lesson time reported here is standardized to 50 minutes.

their students (Figure 2). As the Study Report pointed out, "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)" (Hiebert *et al*, 2003: 109). When we factor in the relatively large class size of the Hong Kong classroom[8], the reticence of the East Asian students is even more striking.

Figure 2: Average Number of Teacher Words to Every One Student Word Per Lesson

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

In the Study, it was found that the lesson time in all the seven countries was dominated by students working on mathematical problems, and thus one of the major units of analysis in the study was the mathematical problems. Different aspects of the characteristics of the problems worked on in the lessons were coded for analysis, and results of some of the analyses are discussed below.

Nature of problem statements

One important characteristic of the mathematical problems is the nature of the problem statements. Three types of problem statements were defined in the Study

---

[8] The average class size of the lessons videotaped in Hong Kong was 37, which is significantly bigger than the class size of other countries in the study (except for Japan for which the class size data was not available) - the average class size in the other five countries ranged from 19 (Switzerland) to 27 (Australia). So a ratio between teacher words and student words of 16 to 1 in Hong Kong is in effect a ratio of nearly 600:1 as far as an individual student in concerned.

based on the kind of mathematical processes implied by the statements. They are using procedures, stating concepts, and making connections (Hiebert *et al*, 2003: 98). Figure 3 below shows the average percentage of problems of each problem statement type in the participating countries[9].
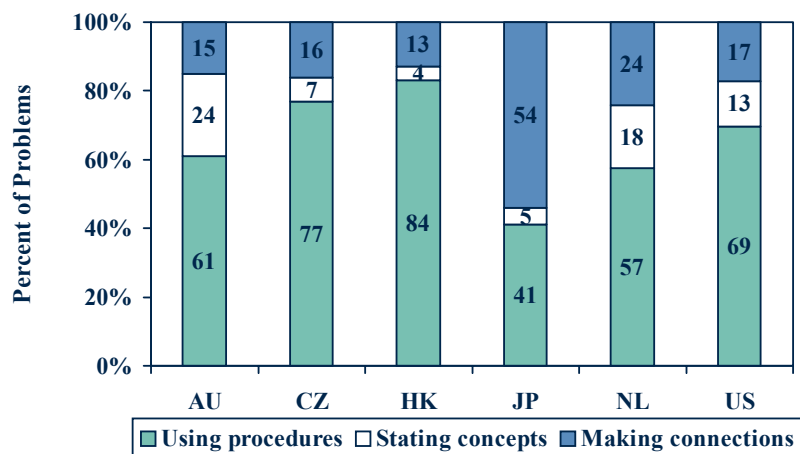

Figure 3: Average Percentage of Problems Per Lesson of Each Problem Statement Type

As can be seen from Figure 3, the problem statements of nearly 85% of the problems worked on in the Hong Kong classrooms suggest that they were typically solved by applying a procedure or a set of procedures. This percentage is highest among all the countries in the Study. Problems with statements that called for mathematical concepts or constructing relationships among mathematical ideas and facts were relatively rare. As the Study Report noted, "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)" (Hiebert *et al*, 2003: 98).

Contexts of the problems

In what contexts were these procedural problems set up when they were presented to the Hong Kong students? Mathematics problems are usually either set up within some real-life contexts or simply presented using mathematical language or symbols (e.g., Solve the equation: $x^2 + 3x - 8 = 0$). Many mathematics educators argue that mathematics problems presented within real-life contexts make mathematics more meaningful and hence more interesting for students.

---

[9] The data from Switzerland was not available since English transcripts were not available for all Swiss lessons.

Figure 4 below shows the average percentage of problems that were set up with a real-life connection compared to those that were presented with mathematical language or symbols only. As can be seen from Figure 4, Hong Kong lessons had most of the problems set up using mathematical language or symbols only, second to Japan. Only 15% of the problems had a real-life connection, and more than 80% of the problems were formulated with mathematical language and symbols only.

Figure 4: Average Percentage of Problems Per Lesson Set Up With a Real Life Connection or
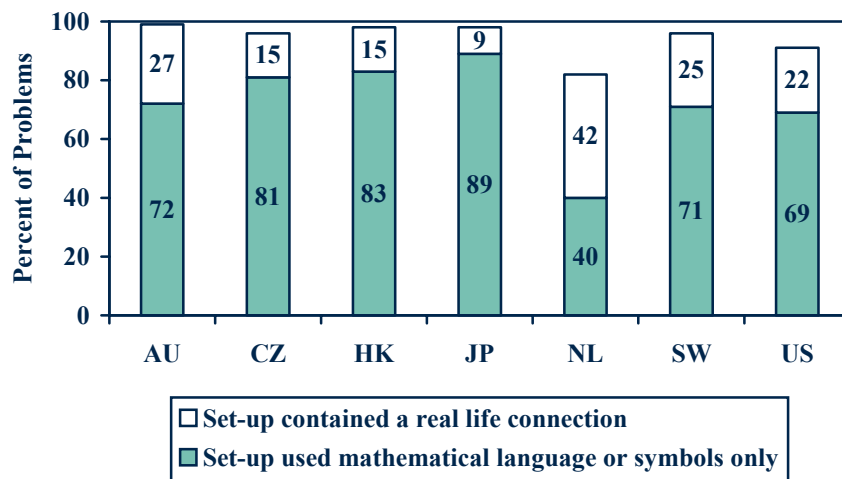
With Mathematical Language or Symbols Only

Choice of solution methods

When Hong Kong students were presented with these procedural problems set up with mathematical language and symbols, how were they expected to deal with the problems? Were they expected to solve the problems with prescribed methods, or were they given a choice and encouraged to solve the problems using different methods? Mathematics educators usually think that to enhance students' problem solving ability, they should be encouraged to solve the same problem with different methods. In the Study, the number of problems worked on in which students had a choice of solution methods was noted, and the results are show in Table 2 below. In Table 2, the left hand column gives the average percentage of problems per lesson in which students had a choice of solution methods, and the right hand column shows the percentage of lessons where there were at least one problem worked on in which students had a choice of solution methods.

Table 2: 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[10]	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

It can be seen from Table 2 that compared with other countries, Hong Kong had the least amount of problems where students were given a choice of solution methods, whether measured by average percentage of problems per lesson or by the percentage of lessons with at least one problem in which students had a choice of solution methods. In only three percent of the problems worked on were students given a choice of solution methods, and such occasions happened in less than 20% of the lessons recorded.

So we can see from the three characteristics of the problems discussed above that the mathematical problems Hong Kong students worked on in their classrooms were mainly problems unrelated to real-life. The statements of the problems suggest that they were typically solved by applying a procedure or a set of procedures rather than calling for mathematical concepts or constructing relationships among mathematical ideas and facts. Furthermore, students were expected to follow prescribed methods in solving these problems instead of being given a choice of solution methods.

Summary

From the results presented above, the instructional practices in the Hong Kong mathematics classroom as portrayed by the analysis of the codes in the Study can be characterized as follows:

---

[10] For the Netherlands, there were too few cases reported and so the data was not shown here because the reporting standard was not met.

Whole-class interaction dominated the lesson time. During the whole-class interaction, the teacher talked most of the time while the students remained relatively reticent. The mathematics problems that students worked on during the lesson were mainly set up using purely mathematical language and symbols, and in contexts unrelated to real-life. These problems were also typically solved by applying a procedure or a set of procedures, following standard methods prescribed by the teacher.

From the viewpoint of most mathematics educators, the picture portrayed above is a mathematics classroom that is not very conducive to quality teaching and learning!

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

As described earlier, one of the qualitative analyses of the data in the Study was performed by an expert panel comprising mathematicians and mathematics educators. Panel members reviewed detailed descriptions of a random sub-sample of the videotaped lessons country-blind and made qualitative judgements about them. In addition to judging the content level of the lessons, the panel also assessed the quality of the mathematics in the lessons along four dimensions: *coherence, presentation, engagement* and *overall quality*. The results of the judgements of the Mathematics Quality Analysis Group are presented below.

More advanced content

The panel made judgement on how advanced the mathematics content in the lessons was, and placed each lesson in the sub-sample into one of five "curricular levels", from elementary (1) to advanced (5). The results of their judgement are shown in Figure 5 below. As can be seen from Figure 5, the panel found the content covered in the Hong Kong (and Czech Republic) classrooms relatively more advanced. The mathematics content of 20% of the lessons was judged to be advanced, while the content in none of the lessons was judged to be elementary. This is in great contrast to other countries in the Study, where three of them (Australia, the Netherlands and the United States) did not have any lessons with content judged to be advanced (and Switzerland had only 5% of the lessons with content judged to be advanced), and the mathematics content in at least 10% of the lessons in four countries was judged to be elementary.
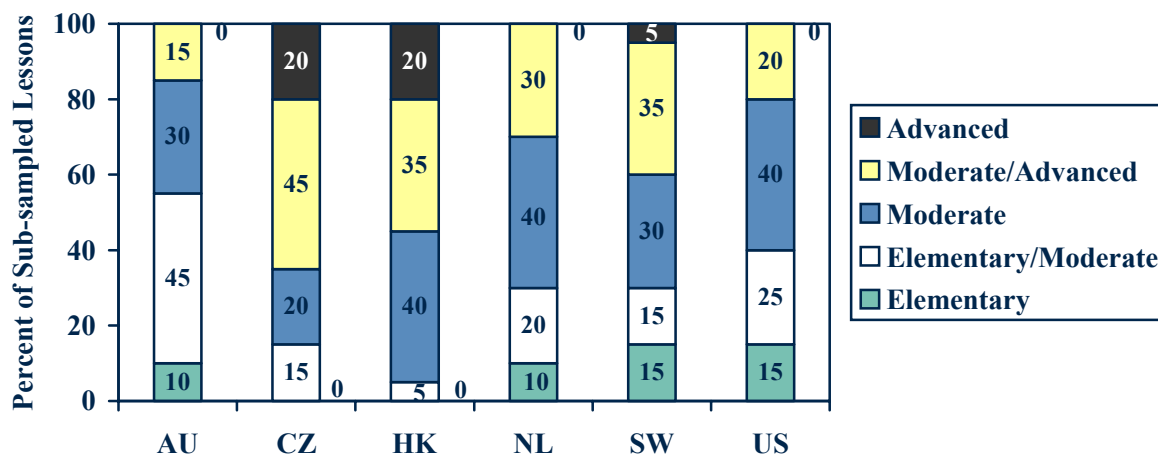
Figure 5: Percentage of lessons in sub-sample at each content level

Lesson more coherent

Coherence was defined by the panel as "the (implicit and explicit) interrelation of all mathematical components of the lesson" (Hiebert *et al*, 2003: 196). As can be seen from Figure 6 below, 90% of the Hong Kong lessons were judged to be thematically coherent, with the remaining 10% moderately thematically coherent. This compares very favorably with the other countries in the Study. For example, in the Czech Republic and the United States, only 30% of the lessons were judged to be thematically coherent.
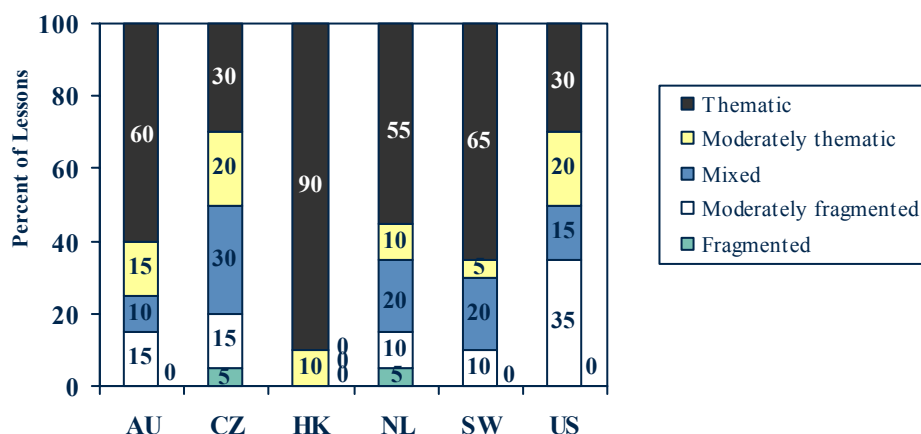

Figure 6: Percentage of Lessons in Sub-sample Rated at Each Level of Coherence

More fully developed presentation

Not only were the Hong Kong lessons judged to be more coherent, their presentation was also found to be more fully developed. Presentation was defined by the panel as "the extent to which the lesson included some development of the mathematical concepts or procedures" (Hiebert *et al*, 2003: 197). 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. The results of the judgment of the panel are show in Figure 7 below. It can be seen from Figure 7 that 20% of the Hong Kong lessons were judged to be "fully developed". This percentage is highest among all the other countries, and is in striking contrast with Australia where none of the lessons were classified as "fully developed". If we take into consideration the category "substantially developed" as well, we can see that three quarters of the lessons in Hong Kong were classified as either 'fully developed' or 'substantially developed'. This figure is three times higher than that for the lessons in the Netherlands.
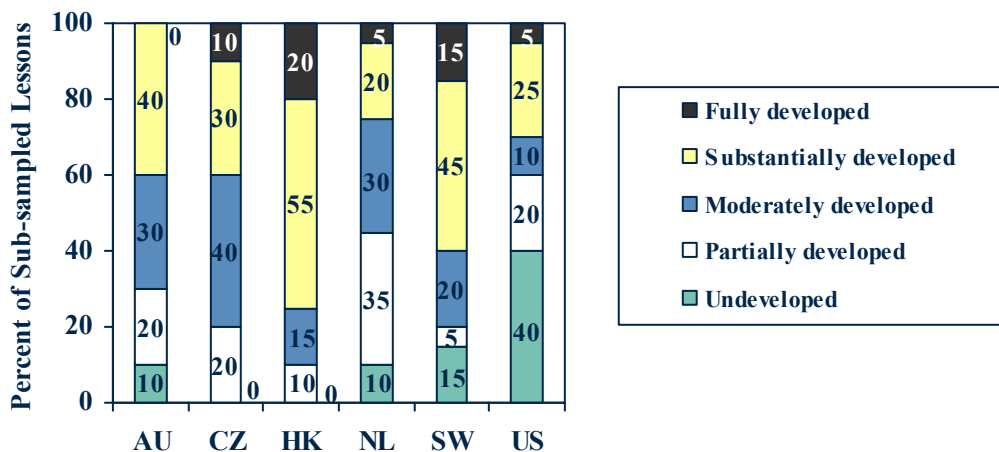

Figure 7: Percentage of Lessons in Sub-sample Rated at Each Level of Presentation

Students more likely to be engaged

As noted previously, the panel did not watch the videotapes (since the exercise was conducted "country-blind") and so could not easily judge whether students were engaged in the lessons or not. From the detailed descriptions of the lessons compiled by the international coding team, the panel made a judgement as to how likely it was that students would be engaged in the lessons. Student engagement

was defined by the panel as "the likelihood that students would be actively engaged in meaningful mathematics during the lesson" (Hiebert *et al*, 2003: 198). A rating of 'very unlikely' (1) indicated a lesson in which students were asked to work on few of the problems in the lesson and those problems did not appear to stimulate reflection on mathematical 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.

As can be seen from Figure 8 below, the panel inferred from the lesson descriptions that students in Hong Kong classrooms were more likely than those elsewhere to be engaged in the lesson. The panel estimated that students in 35% of the Hong Kong lessons were likely to be engaged, whereas in none of the Australian lessons were students likely to be engaged.
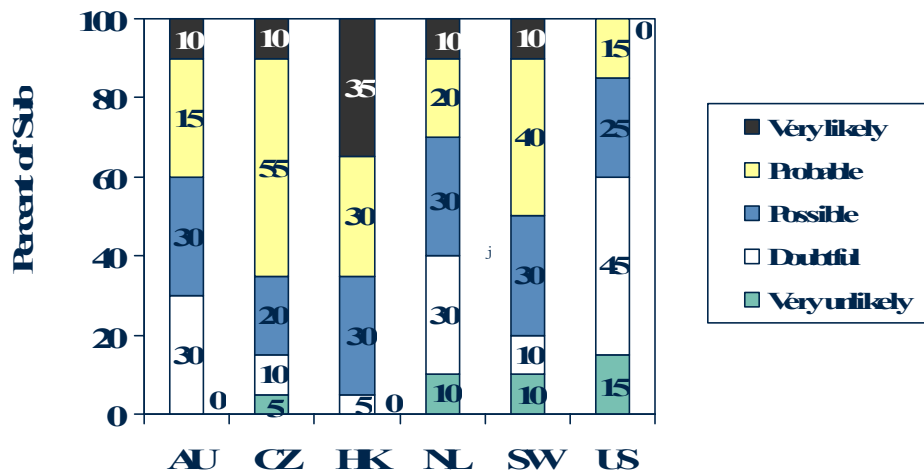

Figure 8: Percentage of Lessons in Sub-sample Rated at Each Level of Student Engagement

Overall quality

Finally, the panel made a judgement on the overall quality of the lessons in terms of "the opportunities that the lesson provided for students to construct important mathematical understandings" (Hiebert *et al*, 2003: 199). Figure 9 below indicates that 30% of the Hong Kong lessons were judged to be of high quality, whereas only 5% of the lessons in Australia and the Netherlands were judged to be so. And in the U.S., none of the lessons were judged to be of high quality. There were also more lessons in Hong Kong than in other countries for which the panel judged the overall quality to be 'high' or 'moderately high'.
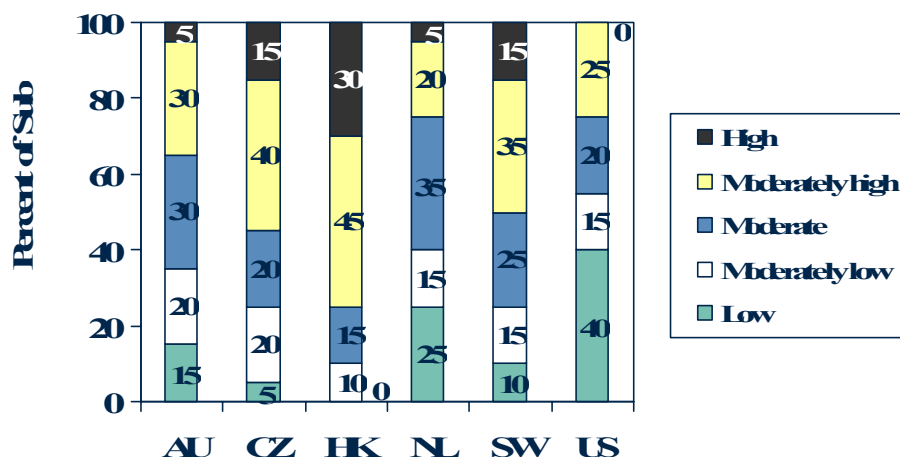
Figure 9: Percentage of Lessons in Sub-sample Rated at Each Level of Overall Quality

Summary

From the results presented above, we can see that the quality of instructional practice in the Hong Kong mathematics classroom was judged by the Mathematics Quality Analysis Group as very high. The mathematics content covered was judged to be relatively more advanced, the lessons were more coherently structured, and the presentation was more fully developed. Given these positive elements of the classrooms, students were expected to be more engaged in the teaching and learning process, and the overall quality of the lessons was judged to be high by the panel.

This presents a picture of instructional practices which is much more positive than that portrayed by the quantitative analysis of the codes.

Discussion

A.   Two different pictures?

From the discussions above, it can be seen that the picture of instructional practices in the Hong Kong classroom as portrayed by the judgement of the Mathematics Quality Analysis Group is in stark contrast to the picture as portrayed by the quantitative analysis of the codes. How do we reconcile the apparent inconsistency between the instructional practices as reflected by the two different analyses of the same data set?

It should be noted that in the first picture, instructional practices were portrayed through objectively coding and summarizing activities that happened in the

classroom, whereas in the second picture, the quality of content was judged by the Mathematics Quality Analysis Group based on their expertise and experience. In the international report of the Study, readers are alerted to the small sample size involved in the qualitative analysis and are urged to be cautious in the interpretation of the results. Readers are warned that 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). Such warning needs to be heeded, for it pertains to the reliability of the analysis results. That is, from a psychometric point of view, the results of the qualitative analysis are deemed to be not very reliable. In addition to the small sample size involved (which is typical of qualitative studies), the very fact that the analysis relied on the judgement of a group of experts means that the results may be "rater-dependent". Given another group of experts with different experience and inclinations, rather different conclusions about the teaching in the Hong Kong classroom may be arrived at, even when the same set of criteria and definitions are followed. In contrast, for the quantitative analysis, since the coding (e.g. number of words spoken by teachers and students) is relatively objective, it is expected that given adequate training, any coder should arrive at more or less the same results.

Since the results of the qualitative analysis are not very reliable statistically, should we discard them and resort only to the reliable quantitative analysis? The quantitative analysis of the TIMSS 1999 Video data, as with all low-inferenced quantitative analysis, has its own limitations as well. Take the number of words spoken by teachers and students in the classrooms as example again. The quantitative analysis of the data computed accurately the number of words spoken by teachers and students in each country, and both the absolute number of words spoken and the ratio between teacher and student words provide relevant information on the kind of interaction that took place in the classrooms concerned. However, every teacher or educator knows too well that the quality of what the teachers and students say in class is far more important that how much they say. But to determine what the teachers and students say are significant or not, the data analysis requires a lot of judgement based on profound experience on the part of the researcher. And in a quantitative study where the emphasis is on low-inferenced data analysis, this is not possible. Hence quantitative analysis may yield results that are highly reliable but not necessarily very meaningful.

Thus, it seems that there is an inherent trade-off between reliability and validity in the analysis of video data. In order to get highly reliable data, we have to restrain from making inferences, and hence we lose out in validity. In order to increase the validity of the analysis of video data, we need to make subjective judgment, with the result that high reliability is difficult to attain. This is rather like Heisenberg's

Principle of Uncertainty in physics[11]. It seems that if we want to get highly reliable and objective information, we have to lose out in the meaningfulness or validity of the data. On the other hand, since qualitative analysis involves the judgement of "experts" based on their experience and expertise, different groups of experts may yield different results. So the information we obtain cannot be very reliable.

Which, then, is the "real" picture of mathematics teaching in Hong Kong? Is the unreliable expert judgment of the Mathematics Quality Analysis Group "real"? Or does the quantitative analysis of the data of the Study fail to reveal the subtlety of the complexity of classroom teaching?

The answer depends on whether you prefer a very reliable description of the activities that happened in the classroom, or whether you can tolerate some lack of reliability and want to learn more about the experts' view on the quality of teaching and learning in the classroom. The crux of the matter is: in determining the quality of teaching, should we rely on objective summary of data, or should we rely on subjective judgment of experts? Perhaps a synthesis of the two gives a picture nearer to the reality.

B.   The Traditional East Asian Culture and the High Achievement of East Asian Students

Given that the results of the quantitative analysis of the Hong Kong data in the Study (which shows that instructional practices in Hong Kong are not very conducive to quality learning) are at least part of the "real" picture in Hong Kong, how can we explain the high achievements of students in Hong Kong and other East Asian countries in international studies of mathematics achievement? Also, do the findings of the Study throw any light on the negative attitudes of East Asian students towards mathematics?

First, the traditional teaching in Hong Kong as revealed by the quantitative analysis of the Study may be explained by the underlying cultural values in East Asia. In a replication of Ma's study (Ma, 1999) in Hong Kong and Korea (Leung and Park, 2002), it was found that although the teachers in the study were in general competent in mathematics, they often deliberately taught in a procedural manner for pedagogical reasons and for the sake of efficiency. They seemed to believe that it would be inefficient or even confusing for school children to be exposed to rich concepts instead of clear and simple procedures. This illustrates

---

[11] The Principle states that the more precisely the position of an object is determined, the less precisely its momentum is known in this instant (see Cassidy, 1992).

very well the pragmatic philosophy in the East Asian culture (Ko, 2001; Shusterman, 2004).

Secondly, the underlying cultural values shared by the East Asian students may also explain both their high achievement and negative attitudes towards mathematics. In the East Asian culture, there is a strong stress on the virtue of humility or modesty. As the author pointed out elsewhere:

> Children from these countries are taught from when they are young that one should not be boastful. This may inhibit students from rating themselves too highly on the question of whether they think they do well in mathematics, and so the scores may represent less than what students are really thinking about themselves. On the other hand, 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)

Given that East Asian students possess such negative attitudes towards mathematics and hold such low self-concept in mathematics, why do they perform so well in international studies of mathematics achievement? Paradoxically, from the standpoint of the East Asian culture, one may argue that this negative correlation between students' confidence in mathematics and their achievement is something to be expected:

> Over-confidence may lower students' incentive to learn further and cause them to put very little effort into their studying, and hence result in low achievement. This is exactly the kind of justification for the stress on humility or modesty in the East Asian culture. The Chinese saying "contentedness leads to loss, humility leads to gain" illustrates the point well. (Leung, 2002: 106)

In addition, the stress in the East Asian culture on diligence and practice may have also contributed to the high achievement of their students (Park and Leung, 2003). Underlying this stress on diligence and practice is the traditional East Asian value that attributes success more to effort than to innate ability (Leung, 2001). The ultimate root of the stress on diligence and practice is the underlying Confucian cultural values which emphasize strongly on the importance of education and a high expectation for students to achieve. Under the influence of this philosophy, learning or studying is considered a serious endeavour, and students are expected

to put in hard work and perseverance in their study. This is reinforced by a long and strong tradition of publication examination, which acts as a further source of motivation for learning. This high expectation on students to achieve provides an important source of motivation for students to learn well and to excel.

C.   *Implications*

Given the methodological complexity of interpreting video data, and the cultural explanation of the high achievement of East Asian students, what lessons can educators from East Asia and elsewhere learn from the results of the Study?

Implications for East Asian countries

First, just examining the results of the quantitative analysis of the Hong Kong data in the Study may prompt us to call for radical changes in instructional practice in the Hong Kong classrooms, and by inference the classrooms in East Asia as well. However results from the qualitative analysis of the data present a different picture. Some readers may tend to embrace the qualitative results since they are more consistent with results of the achievement data (and for readers from East Asia, the results from the qualitative analysis of course look more pleasing!), and dismiss the quantitative results as invalid. But it should be stressed that the quantitative analysis is done using a relatively more objective (at least more objective than the qualitative part of the analysis) methodology and utilizing a larger and more representative data set (compared with the qualitative analysis). So the findings should not be dismissed lightly. A more balanced view of the two sets of results is that they represent two aspects of the same reality. They complement each other in giving a picture closer to the reality of the Hong Kong (and East Asian) classroom.

Seen in this light, findings of the qualitative analysis of the video data should remind educators in East Asia of their strengths in terms of instructional practices in mathematics. In particular, the expectation that students should learn a relatively advanced level of content with an appropriate degree of abstraction ought to be retained. Simply reducing the difficulty of the content in order to make mathematics more accessible to the general student population is an endless retreat. At the same time, teachers from East Asia should treasure their tradition of teaching mathematics in a coherent manner. They should also continue their effort to fully develop their lessons so as to keep their students engaged in the mathematics. Teachers should of course try to make the lesson lively and enjoyable for students through introducing various activities in their classrooms, but the goal should be to induce students to be interested and engaged in the subject matter of mathematics itself rather than in the lively activities per se. These clearly have implications for reforms in curriculum content, teaching methods, the kind of teachers to be recruited, and the kind of teacher education needed for teachers to perform their job.

On the other hand, it should be admitted that dominance of teacher talk may not be the best kind of activities for effective mathematics learning. Also, despite their students' success in international studies of mathematics achievement, educators in East Asia need to ask themselves whether the fact that the majority of the problems student solve are unrelated to real-life is in itself consistent with their ideal of a good mathematics education. The challenge for mathematics educators in East Asia is to promote more student participation in meaningful learning without compromising their strengths in instructional practices as identified above.

Implications for other countries

If instructional practices and the resulting student achievement are so much related to the underlying culture, what are the implications for countries outside East Asia?

First, students' mathematics achievement in international studies should be viewed in conjunction with their attitudes towards mathematics and mathematics learning. Although negative attitudes of students may not necessarily disadvantage their achievements, the negative attitudes themselves should be considered part of the attainment of the curriculum in the countries concerned, and educators should be alarmed by such negative attitudes. Curriculum documents in countries around the world always include enjoyment of study as part of the aims of education, irrespective of the culture of the countries. So a high student achievement in international studies should not relegate efforts to promote students' interest in their study. We don't want students to do well in mathematics while hating it.

Secondly, simple transplant of classroom practices from high achieving countries to low achieving ones would not work. Since teachers and their teaching are so much influenced by the underlying cultural value, 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. What we can learn from another culture through comparative studies is 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 in different cultures, we may then begin to understand the interaction between educational practices and culture, and 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

For many sectors of the community, especially the media, the attention of international comparative studies is usually focused on the relative position of countries in the league tables generated from the studies. For other people, especially the educational policy makers, such international studies sometimes provide an impetus or excuse for educational changes. But very often, such changes are made without a careful consideration of the complex context in different countries within which the achievement and classroom instructions under study are situated.

However, the primary purpose of these international studies is not for countries to compete with each other. Nor should the results of comparative classroom studies be used rashly to justify the classroom practices of the high achieving countries. The significance of such international studies should lie in the rich data set they generate, serving as mirrors for educators to better understand their system. And any changes in educational policies should take into account the rich data set as well as the different cultural values that generate such richness.

References

Beaton, A.E., Mullis, I.V.S., Martin, M.O., Gonzalez, E.J., Kelly, D.L. and Smith, T.A. (1996). *Mathematics Achievement in the Middle School Years,* Centre for the Study of Testing, Evaluation and Educational Policy, Boston College, Boston

Cassidy, D. (1992). "Heisenberg, Uncertainty and the Quantum Revolution", *Scientific American, 266* (May): 106-112.

Hiebert, J., Gallimore, R., Garnier, H., Givvin, K.B., Hollingsworth, H., Jacobs, J., Chui, A.M.Y., Wearne, D., Smith, M., Kersting, N., Manaster, A., Tseng, E., Etterbeek, W., Manaster, C., Gonzales, P. and Stigler, J. (2003). *Teaching Mathematics in Seven Countries. Results From the TIMSS 1999 Video Study,* National Center for Education Statistics, Washington DC.

Ko, S.B. (2001). "China's Pragmatism as a Grand National Development Strategy: Historical legacy and evolution", *Issues & Studies*, 37 (6): 1-28 Nov-Dec.

Leung, F.K.S. (2001). "In Search of an East Asian Identity in Mathematics Education", *Educational Studies in Mathematics*, 47: 35-51.

Leung, F.K.S. (2002). "Behind the High Achievement of East Asian Students"*, Educational Research and Evaluation*, Vol. 8, No. 1, 87-108.

Leung, F.K.S. and Park, K.M. (2002). "Competent Students, Competent Teachers?", *International Journal of Educational Research*, 37(2), 113-129.

Ma, L. (1999). *Knowing and Teaching Elementary Mathematics*. Mahwah, N.J.: Lawrence Erlbaum Associates, Publishers.

Mullis, I.V.S., Martin, M.O., Beaton, A.E., Gonzalez, E.J., Kelly, D.L. and Smith, T.A. (1997). *Mathematics Achievement in the Primary School Years,* Center for the Study of Testing, Evaluation and Educational Policy, Boston College, Boston.

Mullis, I.V.S., Martin, M.O, Gonzalez, E.J., Gregory, K.D., Garden, R.A., O'Connor, K.M., Chrostowski, S.J. and Smith, T.A. (2000). *TIMSS 1999 International Mathematics Report,* International Study Center, Lynch School of Education, Boston College, Boston.

Mullis, I.V.S., Martin, M.O, Gonzalez, E.J. and Chrostowski, S.J. (2004). *TIMSS 2003 International Mathematics Report,* International Study Center, Lynch School of Education, Boston College, Boston.

Organisation for Economic Co-operation and Development (2001). *Knowledge and Skills for Life: First Results from PISA 2000*. Paris: OECD Publications.

Organisation for Economic Co-operation and Development (2003). *Literacy Skills for the World of Tomorrow - Further Results from PISA 2000*. Paris: OECD Publications.

Organisation for Economic Co-operation and Development (2004). *PISA 2003 Report* – in press.

Park, K.M. and Leung, F.K.S. (2003). "Factors Contributing to East Asian Students' High Achievement in Mathematics: The Case of Korea", *The Mathematics Educator*, 1, 7-19.

Shusterman, R. (2004). "Pragmatism and East-Asian thought", *Metaphilosophy*, 35 (1-2): 13-43, Jan.