Eight Grade Students' Mathematics Achievement in TIMSS 2011 Cognitive Domains- A Comparison across Four Southeast Asian Countries

S. Kanageswari Suppiah Shanmugam
SEAMEO RECSAM, Penang, Malaysia

Abstract

This article describes secondary data analyses that explores students' performance in Trends in Mathematics and Science Study (TIMSS) 2011 mathematics items that assess lower order thinking (LOT) and higher order thinking (HOT) in the four participating Southeast Asian countries (Indonesia, Malaysia, Singapore and Thailand). The objectives are to compare the three cognitive domains of *knowing*, *applying* and *reasoning* for these countries, and to study the students' performance in LOT and HOT items. The analyses were done both qualitatively using document analyses to investigate the cognitive processes incorporated in the education system, and quantitatively using IDB Analyzer to determine students' performance. The article reports on detailed students' performance related to the percentage of students at each benchmark and their percentile achievement. The findings reveal that in Indonesia and Thailand, more students were able to answer HOT items, with higher scores than LOT items. For Malaysia, the general student population were able to answer LOT items, with higher scores when compared to HOT items. In Singapore, students at the bottom 25% obtained higher score for LOT items, while the rest performed better on HOT items. However, a higher percentage of students at the weak, intermediate and advanced categories were more able to answer HOT items. The findings seem to suggest that focusing on problem solving may be inadequate. A strategy worth investigating is complementing the infusion of problem solving skills with making mathematical connections to real life situations through non-routine questions. However, more research is required before suggesting any conclusive pedagogical practice.

Keywords: cognitive dimension, higher order thinking, lower order thinking, TIMSS 2011 Mathematics

Introduction

Trends in Mathematics and Science Study (TIMSS) grade eight mathematics items are organised around two dimensions, which are content and cognitive. While the content dimension describes the mathematical concepts that are being assessed, the cognitive dimension characterises the thinking processes that underlie each item (Mullis, Martin, Ruddock, Sullivan, & Preuschoff, 2009). The cognitive dimension, which is the focus of this paper, identifies the cognitive skills that students rely on when solving each item. These descriptions are important as they provide evidence of the type of thinking that is involved when students are in the process of applying their mathematical knowledge (National Council of Teachers of Mathematics, 2000).

While solving the items, students draw on a range of cognitive skills necessary in accessing mathematical knowledge that lead to a successful solution. They tap into different types of thinking, which among others include critical thinking, creative thinking, deductive thinking, inductive thinking, rational thinking, analogical thinking, metaphor, metacognition,

making inferences, making generalisation, making conclusion, judging idea, making predictions, solving problems, analysing, proposing solutions, and comparing and hypothesising. These wide range of cognitive skills are subsumed under two large constructs of higher order thinking (HOT) and lower order thinking (LOT) (Rajendran, 2010).

This paper intends to explore students' performance in TIMSS items that assess LOT and HOT, and is organised into several sections that firstly describes items in TIMSS 2011 that assess LOT and HOT before examining students' detailed performance (percentage of students at each international benchmark and percentile score) in each category. The cognitive processes that are incorporated in the education system of the four Southeast Asian countries (Indonesia, Malaysia, Singapore and Thailand) are examined by analysing the documents that outline their Mathematics Curriculum. The rationale of confining the analyses to these four countries is mainly because these are the only four countries in the Southeast Asia region that participated in TIMSS 2011, in addition to the sharing commonalities like the western colonialism that affected the historical education background, race, religion and languages spoken.

Lower Order Thinking and Higher Order Thinking

Teaching students higher order thinking (HOT) skills and assessing them with items that measure their higher order thinking are two important educational agendas that are popularly gaining global attention and recognition. The reasons behind it are many, but the association between HOT and students' higher achievement, and the relevance of HOT to the present 21st century learning (Thompson, 2011) are prevalent justifications.

While the significance of LOT and HOT are acknowledged in the education community, the definitions by far are varied and different for each one of them, yet commonalities string the different descriptions. Maier (1933) referred to LOT as learned behaviour or reproductive thinking and HOT as reasoning or productive behaviour. These terms appropriately address the distinctive difference between them; LOT occurs when the solution involves reproducing a previously learned algorithm or a familiar task, while HOT occurs while productively solving unfamiliar tasks using past experiences that have not been previously associated. Newman (1990) too asserts that LOT requires the application of previously learned information, making the task a routine one, while HOT occurs when students manipulate information to solve a non-routine challenging task. He also pointed out that tasks that require LOT or HOT are dependent on the students' intellectual history.

According to Rajendran (2010), HOT is the amplified use of the mind that occurs when faced with new and non-routine challenges that requires interpretation, analyses or manipulation of information. LOT, on the other hand, involves the restricted use of the mind that allows recall of information or application of previously learned knowledge on routine problems.

The commonalities in these descriptions are that HOT is associated with solving non-routine questions using non-rehearsed algorithm and in contrast, LOT occurs when students solve familiar 'text book' problems or routine tasks using well-known algorithms.

TIMSS Cognitive Dimensions

The TIMSS cognitive dimension has three domains that are *knowing, applying*, and *reasoning*, which receives different emphases. The *knowing* domain receives 35 % of the total testing, while the *applying* domain receives the most emphasis, with a testing time of 40%. The *reasoning* domain is allotted the least testing time of 20% (Mullis et al., 2009).

The *knowing* domain assesses students on mathematical facts, concepts, and procedures. Mathematical facts include the knowledge of basic mathematics register and number properties. Mathematical concepts focus on the knowledge of making connections between the bodies of learnt language, making judgment about mathematical statement and the methods employed, and making mathematical representations. Mathematical procedures encompass a range of computational procedures and tools used to solve routine problems, which require students to recall pre-taught set of actions and execute them. They connect the basic knowledge to using the knowledge for solving everyday problems (Mullis et al., 2009).

On the surface, the *knowing* domain seems to be exploring simple and fundamental cognitive skills of recalling of basic facts and conventions of numbers. A thorough exploration however, reveals that a good grasp of these pre-requisites is the key for students to progress to the more advanced skills related to problem solving and abstract mathematical thinking. Without a good knowledge of mathematical facts, concepts and procedures, students are challenged when they climb up the hierarchical structure of mathematics. The lists of behaviours that characterise the *knowing* domain are recall, recognise, compute, retrieve, measure and classify or order (Mullis et al., 2009).

The *applying* domain tests students' understanding of the mathematical knowledge to solve routine problems. Specifically, they display the ability to select, represent, model, implement, and solve routine problems. These routine problems with varying difficulty level are exercises that students have practiced during classroom instruction, which may range from real-life contexts to authentic mathematics questions. In order to solve these problems, students apply the mathematical facts, concepts, and procedures that they are familiar with to create mathematical representations. Creating mathematical representations is crucial as it gradually develops students' mathematical thinking and communication, which are necessary in the next cognitive level of *reasoning* (Mullis et al., 2009).

The *reasoning* domain is an advanced cognitive process that stretches beyond solving routine problems of applying mathematical knowledge to the familiar 'text-book' problems. *Reasoning* taps into students' logical thinking of inductive and deductive reasoning to determine solutions to non-routine problems, which students achieve by studying patterns. Just like in the *applying* domain, the problems can be real-life situations or authentic mathematics items. However, the *reasoning* domain requires students to solve non-routine, complex, and multi-step problems, which demand cognitive skills that are beyond those required to solve routine problems. The challenge faced in solving these non-routine problems is mastering the mathematical knowledge and applying that mathematical knowledge may just not be adequate. The novelty of the context created by the items compels students to reason out their answers, breaking away from the norm of working out rehearsed or routine steps to a solution. Specifically, it describes students' behaviour of being able to analyse, generalise, synthesise, justify, and solve non-routine problems (Mullis et al., 2009).

Based on the descriptions of these three cognitive domains, and the definitions of LOT and HOT used within the context of this paper, the TIMSS cognitive domains of *knowing* and *applying* appear to be assessing LOT, while *reasoning* is related to HOT. To further elucidate, while it is clear that the *knowing* and *reasoning* domains are assessing LOT and HOT respectively, there seems to be a shadow of doubt in the *applying* domain. This is because the nature of item, specifically the familiarity of the item distinguishes LOT from HOT. However, based on the description provided in TIMSS framework, the *applying* domain assesses items that are routine or familiar and as such, within the context of this paper, the items in the *applying* domain is classified as assessing LOT. As such, within the scope of this paper, the items in the *knowing* and *applying* domains assess students' LOT, while items in the *reasoning* domain assess students' HOT.

TIMSS Grade Eight Mathematics Items

As the extent of using the mind ('stretched' verses 'restricted') is a critical distinction between HOT and LOT, the type of item used to assess students is an important factor to consider. Even though there are claims that LOT and HOT can be assessed by any item type, there are still reservations to the efficiency of multiple-choice items assessing HOT as it fails to engage students' disposition. The contributing factors are mainly due to its unique format of 'one right answer', which restrains students from justifying or reasoning their answers. On the contrary, constructed-response items invokes information like students' background knowledge and the strategies they employ for solving questions (Quellmalz, 1985).

The TIMSS 2011 items for grade eight consist of both multiple-choice and constructed-response items, and are organised around the three cognitive domains of *knowing, applying* and *reasoning*. A total of 217 items, with score points of 232 were rotated around 14 booklets. Each booklet contained two blocks of Mathematics items, with each block constituting 12 to 18 items. While the score point for multiple-choice item was fixed at one point, the score points for constructed response items ranged from one to two points. As such, a block of items carried a maximum of 18 points (Mullis, Martin, Foy, & Arora, 2012).

Table 1
Distribution of Assessment Items by Cognitive Domains and Item Type

Item Type	Knowing	Applying	Reasoning
	Num of items (%)		
Multiple-choice Items	53 (66.25)	47 (55.29)	18 (34.62)
Constructed-response Items	27 (33.75)	38 (44.71)	34 (65.38)
Total	80 (100.00)	85 (100.00)	52 (100.00)

The data on the number of multiple-choice items and constructed-response items for each of the cognitive domains (available from *TIMSS 2011 International Results in Mathematics*) was used to determine the percentage of items in each of the domains. From Table 1, it can be inferred that the *applying* domain receives the most attention, with 85 items in this category, while the fewest number of items is in the *reasoning* domain with only 52 items. The highest number of multiple-choice items (53) is from the *knowing* domain and the highest number of constructed-response items (38) is from the *applying* domain. Fine grain analysis reveals that the multiple-choice items were hugely used to assess LOT of *knowing* (66.25%) and *applying* (55.29%), while the least was used to assess HOT in the *reasoning* domain (34.62). This is

probably due to the nature of the multiple-choice items that limit the demands of the cognitive process to only recalling, recognising and applying the principles of mathematics on a rehearsed item. On the other hand, the smallest percentage of constructed-response items was used to assess the *knowing* domain (33.75%) and followed by the *applying* domain (44.71%). As suggested by the nature of constructed-response items in assessing the extended use of the mind, constructed-response items constituted the highest percentage of 65.38% in assessing HOT in the *reasoning* domain. While a bigger percentage of items from the cognitive domains of *knowing* and *applying* were multiple-choice items, the *reasoning* domain constituted a bigger percentage of constructed-response items.Mathematical Ability of Students in the Four Countries

Mathematical Ability of Students in the Four Countries

This section provides an overview on students' overall mathematical ability in answering TIMSS items that assessed LOT and HOT skills. In order to gain a general understanding on their acquired levels of mathematical ability and skills, the average scale scores, average percent correct and percentage of students at international benchmarks for the three cognitive domains were examined.

In this section, Table 2 exhibits the average scale score each of the three cognitive domains in the four countries. Since the knowing domain tests simple and basic cognitive skills of recall, which serves as the pre-requisite for the more challenging domains of applying and reasoning, the assumption is that students' achievement will be higher in the knowing domain than in the other two domains. As can be seen in Table 2, this is not the case for all the countries. The assumption holds true for Malaysia and Singapore whose subscale scores in the knowing domain are significantly higher than the overall mathematics score, and also the highest when compared to the other two hierarchically higher domains of applying and reasoning. In addition, the subscale scores consistently dropped when the demands for the cognitive process increased as indicated by the consistently decreasing subscale scores from the knowing domain to the applying and reasoning domains. Another observation is that Malaysia also recorded a significantly lower subscale score in the reasoning domain that the overall mathematics score.

Table 2
Students' Mathematics Achievement in the Cognitive Domains

Country/	Overall Mathematics Average Scale Score (s.e.)	Knowing Average Scale Score (s.e.)	Applying Average Scale Score (s.e.)	Reasoning Average Scale Score (s.e.)
Indonesia	386 (4.3)	378 (4.8) ↓	384 (4.7)	388 (3.8)
Malaysia	440 (5.4)	444 (5.7) ↑	439 (5.2)	426 (5.5) ↓
Singapore	611 (3.8)	617 (3.8) ↑	613 (3.9) ↑	604 (4.3)
Thailand	427 (4.3)	423 (4.7) ↓	428 (4.1)	429 (4.3)

Note. ↑ Subscale score is significantly higher than the overall mathematics score
↓ Subscale score is significantly lower than the overall mathematics score
From "TIMSS 2011 International Results in Mathematics" by Mullis, Martin, Foy, and Arora, (2012, p. 150).

Indonesia and Thailand, on the contrary, performed significantly lower in the *knowing* domain than in the overall mathematics score and also recorded the lowest subscale score in the *knowing* domain when compared to the other two domains of *applying* and *reasoning*. The subscale scores also decreased as the demand for the cognitive process increased as the highest subscales score was for the HOT items in the *reasoning* domain.

The descriptive data as exhibited in Table 2 inform that Malaysia and Singapore students performed better on LOT items from the *knowing* and *applying* domains than HOT items in the *reasoning* domain. But, Indonesia and Thailand performed better on HOT items from the *reasoning* domain than on LOT items from the *knowing* and *applying* domains.

Table 3
Average Percent Correct in the Cognitive Domains

Countries	Overall Mathematics	Knowing	Applying	Reasoning
Indonesia	24 (0.6)	31 (0.7)	23(0.6)	17(0.4)
Malaysia	34 (1.0)	44(1.2)	33(1.0)	23(0.9)
Singapore	73 (0.9)	82(0.8)	73(1.0)	62(1.1)
Thailand	31 (0.9)	38(1.0)	30(0.8)	22(0.8)

Note. From "TIMSS 2011 International Results in Mathematics" by Mullis, Martin, Foy, and Arora, (2012, p. 462).

Table 3 displays the average percent correct (AVC) for the *knowing, applying* and *reasoning* domains. For all the four countries, the AVC in the *knowing* domain is higher than in the overall mathematics and also the highest when compared to the other two cognitive domains. As such, it can be concluded that for all the four countries the AVC declined as the cognitive demands increased from *knowing* to *reasoning* and that, the students tend to answer correctly more items in the *knowing* domain, followed by the *applying* domain and the

reasoning domain. In addition, items in the *reasoning* domain is the least answered correctly by the students in all these four countries. This brings to the conclusion that students in these four countries are generally less able to answer correctly HOT items than LOT items.

In addition, the four TIMSS 2011 international benchmarks: advanced (625 score points or above), high (550 or above), intermediate (475 or above) and low (scoring 400 or above) highlight the cognitive skills that are associated at each (Mullis, Martin, Foy, & Arora, 2012, p. 114). From these cognitive skills, it can be summarised that students at the advanced benchmark exhibit the highest level of mathematical ability, which is to reason, and draw and justify conclusions, while students at the high benchmark project a level below of that to the advanced group, which is to analyse, apply and solve question that involve multi-step problems. Students at the intermediate benchmark are able to apply the mathematical knowledge to solve problems, while students at the low benchmark demonstrate the minimal cognitive level of recalling or retrieving some mathematical knowledge to compute simple operations. The students below the 'low' category are below the minimal expected level of 'having some knowledge of whole numbers, decimal and operations'. In conclusion, the students at the advanced and high benchmarks are able to reason, while students at the intermediate benchmark demonstrate the ability to apply, while students at the low benchmark are at the cognitive level of knowing mathematical knowledge. This also implies that students below the low benchmark have not reached the cognitive domain of *knowing* as they lacked the minimal level of possessing some knowledge of number and operations.

In order to determine the percentage of students demonstrating these varied cognitive skills at each benchmark, the data on cumulative percentage for students 'at or above' the score representing each benchmark, (accessible from p. 114 of the *TIMSS 2011 International Results in Mathematics*) was converted to percentage of students at each benchmark by using subtraction. For example, Student % at Low benchmark = Student % 'at 400 or above' – Student % 'at 475 or above'. These obtained values are exhibited in Table 4.

Table 4
Percentage of Students at the Advanced, High, Intermediate, Low and Below the Low Benchmarks

Country	Student % below the Low Benchmark	Student % at the Low Benchmark	Student % at the Intermediate Benchmark	Student % at the High Benchmark	Student % at the Advanced Benchmark
Indonesia	57.14	28.29	12.12	2.27	0.18
Malaysia	34.58	29.05	24.03	10.78	1.56
Singapore	1.10	6.42	14.66	30.18	47.64
Thailand	38.30	33.56	20.42	6.18	1.54

Note. Adapted from "TIMSS 2011 International Results in Mathematics" by Mullis, Martin, Foy, and Arora, (2012, p. 114)."

From Table 4, it can be inferred that generally more than half of the Indonesian student population had not acquired the basic knowledge of Mathematics, while a very small percentage of 2.45% at the advanced and high benchmarks demonstrated the ability to analyse, apply, reason, make and justify conclusion, and make generalisation. The rest (40.41%) seemed to exhibit the mathematical ability of *knowing* and *applying*, such as solving, applying and reading tables. For Malaysia, approximately 35% of students were not able to answer items that demonstrated they had some knowledge on numbers and operations, while students at the advanced and high benchmarks composed of 12 %, and exhibited the ability to reason. A slightly bigger group of 47% at the low and intermediate benchmarks exhibited the ability of knowing mathematical facts and applying them to solve questions. Singapore, on the other hand, has a bigger group of students at the advanced and high benchmarks, who were able to reason and a minute 1.1% of students who did not know basic mathematical knowledge. Around 38% of Thai students could not answer items that indicated that they possessed basic knowledge of mathematics, while approximately 8% of them were able to reason and had reached the advanced and high benchmarks

The discussions so far were focussed on students' general performance for each cognitive domain, without considering differences between matched to their mathematical ability. The next section will examine performance of students with similar mathematical ability by comparing their achievement in the three cognitive domains to the corresponding ability level.

Mathematical Ability of Students in the Four Countries

This section provides an overview on students' overall mathematical ability in answering TIMSS items that assessed LOT and HOT skills. In order to gain a general understanding on their acquired levels of mathematical ability and skills, the average scale scores, average percent correct and percentage of students at international benchmarks for the three cognitive domains were examined.

In this section Table 2 exhibits the average scale score each of the three cognitive domains in the four countries. Since the *knowing* domain tests simple and basic cognitive skills of recall, which serves as the pre-requisite for the more challenging domains of *applying* and *reasoning*, the assumption is that students' achievement will be higher in the *knowing* domain than in the other two domains. As can be seen in Table 2, this is not the case for all the countries. The assumption holds true for Malaysia and Singapore whose subscale scores in the *knowing* domain are significantly higher than the overall mathematics score, and also the highest when compared to the other two hierarchically higher domains of *applying* and *reasoning*. In addition, the subscale scores consistently dropped when the demands for the cognitive process increased as indicated by the consistently decreasing subscale scores from the *knowing* domain to the *applying* and *reasoning* domains. Another observation is that Malaysia also recorded a significantly lower subscale score in the *reasoning* domain that the overall mathematics score.

Table 2
Students' Mathematics Achievement in the Cognitive Domains

Country/	Overall Mathematics Average Scale Score (s.e.)	Knowing Average Scale Score (s.e.)	Applying Average Scale Score (s.e.)	Reasoning Average Scale Score (s.e.)
Indonesia	386 (4.3)	378 (4.8) ↓	384 (4.7)	388 (3.8)
Malaysia	440 (5.4)	444 (5.7) ↑	439 (5.2)	426 (5.5) ↓
Singapore	611 (3.8)	617 (3.8) ↑	613 (3.9) ↑	604 (4.3)
Thailand	427 (4.3)	423 (4.7) ↓	428 (4.1)	429 (4.3)

Note. ↑ Subscale score is significantly higher than the overall mathematics score
  ↓ Subscale score is significantly lower than the overall mathematics score
From "TIMSS 2011 International Results in Mathematics" by Mullis, Martin, Foy, and Arora, (2012, p. 150).

Indonesia and Thailand, on the contrary, performed significantly lower in the *knowing* domain than in the overall mathematics score and also recorded the lowest subscale score in the *knowing* domain when compared to the other two domains of *applying* and *reasoning*. The subscale scores also decreased as the demand for the cognitive process increased as the highest subscales score was for the HOT items in the *reasoning* domain.

The descriptive data as exhibited in Table 2 inform that Malaysia and Singapore students performed better on LOT items from the *knowing* and *applying* domains than HOT items in the *reasoning* domain. But, Indonesia and Thailand performed better on HOT items from the *reasoning* domain than on LOT items from the *knowing* and *applying* domains.

Table 3
Average Percent Correct in the Cognitive Domains

Countries	Overall Mathematics	Knowing	Applying	Reasoning
Indonesia	24 (0.6)	31 (0.7)	23(0.6)	17(0.4)
Malaysia	34 (1.0)	44(1.2)	33(1.0)	23(0.9)
Singapore	73 (0.9)	82(0.8)	73(1.0)	62(1.1)
Thailand	31 (0.9)	38(1.0)	30(0.8)	22(0.8)

Note. From "TIMSS 2011 International Results in Mathematics" by Mullis, Martin, Foy, and Arora, (2012, p. 462).

Table 3 displays the average percent correct (AVC) for the *knowing, applying* and *reasoning* domains. For all the four countries, the AVC in the *knowing* domain is higher than in the overall mathematics and also the highest when compared to the other two cognitive domains. As such, it can be concluded that for all the four countries the AVC declined as the cognitive demands increased from *knowing* to *reasoning* and that, the students tend to answer correctly more items in the *knowing* domain, followed by the *applying* domain and the *reasoning* domain. In addition, items in the *reasoning* domain is the least answered correctly by the students in all these four countries. This brings to the conclusion that students in these four countries are generally less able to answer correctly HOT items than LOT items.

In addition, the four TIMSS 2011 international benchmarks: advanced (625 score points or above), high (550 or above), intermediate (475 or above) and low (scoring 400 or above) highlight the cognitive skills that are associated at each (Mullis, Martin, Foy, & Arora, 2012, p. 114). From these cognitive skills, it can be summarised that students at the advanced benchmark exhibit the highest level of mathematical ability, which is to reason, and draw and justify conclusions, while students at the high benchmark project a level below of that to the advanced group, which is to analyse, apply and solve question that involve multi-step

problems. Students at the intermediate benchmark are able to apply the mathematical knowledge to solve problems, while students at the low benchmark demonstrate the minimal cognitive level of recalling or retrieving some mathematical knowledge to compute simple operations. The students below the 'low' category are below the minimal expected level of 'having some knowledge of whole numbers, decimal and operations'. In conclusion, the students at the advanced and high benchmarks are able to reason, while students at the intermediate benchmark demonstrate the ability to apply, while students at the low benchmark are at the cognitive level of knowing mathematical knowledge. This also implies that students below the low benchmark have not reached the cognitive domain of *knowing* as they lacked the minimal level of possessing some knowledge of number and operations.

In order to determine the percentage of students demonstrating these varied cognitive skills at each benchmark, the data on cumulative percentage for students 'at or above' the score representing each benchmark, (accessible from p. 114 of the *TIMSS 2011 International Results in Mathematics*) was converted to percentage of students at each benchmark by using subtraction. For example, Student % at Low benchmark = Student % 'at 400 or above' – Student % 'at 475 or above'. These obtained values are exhibited in Table 4.

Table 4
Percentage of Students at the Advanced, High, Intermediate, Low and Below the Low Benchmarks

Country	Student % below the Low Benchmark	Student % at the Low Benchmark	Student % at the Intermediate Benchmark	Student % at the High Benchmark	Student % at the Advanced Benchmark
Indonesia	57.14	28.29	12.12	2.27	0.18
Malaysia	34.58	29.05	24.03	10.78	1.56
Singapore	1.10	6.42	14.66	30.18	47.64
Thailand	38.30	33.56	20.42	6.18	1.54

Note. Adapted from "TIMSS 2011 International Results in Mathematics" by Mullis, Martin, Foy, and Arora, (2012, p. 114)."

From Table 4, it can be inferred that generally more than half of the Indonesian student population had not acquired the basic knowledge of Mathematics, while a very small percentage of 2.45% at the advanced and high benchmarks demonstrated the ability to analyse, apply, reason, make and justify conclusion, and make generalisation. The rest (40.41%) seemed to exhibit the mathematical ability of *knowing* and *applying*, such as solving, applying and reading tables. For Malaysia, approximately 35% of students were not

able to answer items that demonstrated they had some knowledge on numbers and operations, while students at the advanced and high benchmarks composed of 12 %, and exhibited the ability to reason. A slightly bigger group of 47% at the low and intermediate benchmarks exhibited the ability of knowing mathematical facts and applying them to solve questions. Singapore, on the other hand, has a bigger group of students at the advanced and high benchmarks, who were able to reason and a minute 1.1% of students who did not know basic mathematical knowledge. Around 38% of Thai students could not answer items that indicated that they possessed basic knowledge of mathematics, while approximately 8% of them were able to reason and had reached the advanced and high benchmarks

The discussions so far were focussed on students' general performance for each cognitive domain, without considering differences between matched to their mathematical ability. The next section will examine performance of students with similar mathematical ability by comparing their achievement in the three cognitive domains to the corresponding ability level.

Research Aim and Objectives

This research is focussed on studying the cognitive domains in the four countries as an attempt to detect any pattern of similarities or differences that can explain achievement in these four countries. The research objectives are, in these four countries

1. What are the similarities and differences for the three cognitive domains?
2. How have the students been performing for these cognitive domains?
3. How have the students been performing in the LOT and HOT items?

Methodology

The main type of data analyses is curriculum content analysis in the four Southeast Asia countries that participated in TIMSS 2011. Data for the curriculum content analysis was based on the information provided by four countries for the TIMSS assessment and is documented in *TIMSS 2011 Enclopedia (*Mullis et al., 2012a, 2012b). Secondary data analyses were conducted using the IEA IDB Analyzer (version 3.0.47) which is available from The International Association for the Evaluation of Educational Achievement (IEA) website. IDB Analyzer creates SPSS code to conduct statistical analyses without compromising the complex sample structure of the huge database. It allows data from different countries to be combined for cross-country analysis, which range from computing means or percentages for variable to regression analyses (http://www.iea.nl/data.html).

Data Analyses

In an attempt to answer the first research question on the similarities and differences of the mathematical processes that the students at the end of Grade Eight are expected to undergo, the information from the *TIMSS 2011 Enclopedia* prepared by the respective countries on the mathematics curriculum for the four countries were compared. In answering the second and third research questions, released data from TIMSS 2011 was analysed using IDB Analyzer. Benchmark and percentile analyses were conducted for each of the three cognitive domains for the four countries. Percentage of students for each category was analysed using IDB Analyzer so that students' performance could be compared for each international benchmark for the countries across the three cognitive domains. The cumulative percentage that was obtained 'at or above' each benchmark was further recomputed using simple subtractions to obtain the percentage of students at the respective benchmark, just like the earlier calculations for Table 4. For example,

Student % at Low Benchmark = Student % 'at or above' - Student % 'at or above'

(labelled as Low)   Low Benchmark   Intermediate Benchmark.

The percentage of students, who scored below the 'low international benchmark' was relabelled as the 'weak' category and it was calculated by subtracting the '*Student % at Low Benchmark*' from 100%. This percentage represents the group of students who were not able to achieve the minimum level of the low benchmark and will be discussed in relation to the percentage of students who are able to answer LOT and HOT items.

Results

Research Question 1

In order to study the mathematical processes that were expected to occur among students at the end of Grade Eight, the three cognitive domains were compared among Indonesia, Malaysia, Singapore and Thailand. The similarity is that the Mathematics curriculum in all these four countries incorporated the mathematical processes covered in all the three cognitive domains of *knowing*, *applying* and *reasoning*, while an obvious difference is that these domains received different emphases.

Indonesia. The mathematics process covered by the students at the end of Grade Eight Mathematics Curriculum includes the mathematical skills and process in the domains of *knowing* (understanding operations, finding missing terms, identifying angles or congruent

triangles), *applying* (solving, calculating) and *reasoning* (interpreting data set, drawing conclusions, making predictions, judging probability of event) (Rahmawati, 2012). The Indonesian students who sat for TIMSS 2011 are the product of the new teaching styles for mathematics, known as *Pendidikan Matematika Realistik Indonesia* (PMRI), which is an Indonesian adaptation of Realistic Mathematics Education with roots back to the Netherlands. Indonesia adopted PMRI partly because the findings from TIMSS 1999 Video Study identified RME as one of the factors that contributed to the Dutch students' increased achievement in Mathematics (Sembiring, Hadi & Dolk, 2008). PMRI was implemented in 2000 (Reba'i, 2013).

PMRI emphasises problem solving that encourages students to be active thinkers and advocates the problem solving culture in the mathematics classroom. The mathematics lessons emphasise contextual problems that are relevant and meaningful to students' daily life. Since the implementation of PMRI, the transformation that occurred in the teaching techniques required students to move beyond answering to understanding other students' answers, providing reasons to accept or disapprove those answers, explore other possible solutions and reflect on the day's lesson (Sembiring, Hadi & Dolk, 2008). As such, it can be concluded that the mathematical process in the Indonesia mathematics education involve the three cognitive domains of *knowing*, *applying* and *reasoning*. However, just like any educational reformations, PMRI was slowly being accepted by the teaching community. This is because the traditional teaching of mathematics being a 'fixed' subject and students as passive learners had long been practiced in Indonesia classrooms (Sembiring, 2010).

Malaysia. The Malaysian Mathematics Curriculum for secondary students carries the objective of producing students who are able to think mathematically, rationally and logically, apply mathematical knowledge to solve problems effectively, make decisions and communicate mathematically. Specifically, the curriculum is designed to ensure that students understand the mathematical facts and operations, solve problems, make decisions, and communicate mathematically in oral and written forms. In ensuring that students acquire the mathematical skills, the curriculum clearly states that teaching and learning process in the classroom must incorporate the developments of problem solving skills (in accordance to Polya's Problem Solving Techniques), and logical, systematic and creative thinking skills (Malaysia. Ministry of Education, 2004b).

The curriculum clearly states that the emphases are on problem solving and mathematics communication, which are incorporated in the three learning areas of Numbers, Shapes and Space, and Relationship. The mathematical skills that are stressed include the cognitive domains of *knowing* (perform computation, understand basic measurement, determining image, identifying geometric properties), *applying* (solving problems) and *reasoning* (interpreting data) (Muhammad Zaini Mohd Zaini & Dewani Goloi, 2012). In addition, the secondary mathematics also focuses on mathematical connections and representations as stated in the Curriculum Specifications for Mathematics Form 2. Mathematics connections allows students to see Mathematics as an integrated whole rather than fragmented ideas, and to apply and reason in order to connect the mathematics learnt in the classroom to the mathematics in real-life situations. Representation allows students to analyse a problem before interpreting them (Malaysia. Ministry of Education, 2004a).

In comparison, just like Indonesia, Malaysian mathematics education covers the cognitive processes in the three cognitive domains of *knowing*, *applying* and *reasoning*, and stresses incorporating problems that are real and relevant to students' daily lives. However, the curriculum clearly states the importance of developing mathematical connections, representation and mathematical communication among students.

Singapore. The Singapore mathematics curriculum, is aimed at developing students' mathematical ability, with a clear focus on developing their problem solving ability. As stipulated in the Singapore Mathematics Curriculum Framework, there are five components directed towards developing students' problem solving ability, which are *attitudes*, *metacognition*, *process, concepts* and *skills*. While the component on *attitudes* is self-explanatory, *metacognition* addresses monitoring students' own thinking, *concepts* refers to numerical, algebraic, geometrical, statistical, probabilistic and analytical concepts, and *skills* relate to numerical calculation, algebraic manipulation, spatial visualisation, data analysis, measurement, use of mathematical tools and estimation. The component on *process* stresses the cognitive skills like reasoning, communication and connections, thinking skills and heuristics, and applications and modelling (Chin, 2012).

The Secondary Mathematics Syllabus states that, mathematical problem solving is the hub of mathematics learning and that "it involves the acquisition and application of mathematics concepts and skills in a wide range of situations, including non-routine, open-ended and real-world problems" (Singapore. Ministry of Education, 2006, p. 2). In studying

the similarities among the mathematics curriculum of Indonesia, Malaysia and Singapore, the focus is on problem solving skills and using problems related to students' daily lives. In contrast, the Singapore mathematics curriculum implicitly states *modelling* and *heuristics* as mathematical processes that need to be incorporated, and clearly adds an emphasis on using non-routine open ended problems as an attempt to expose students to a wider range of situations.

Thailand. The mathematics curriculum in Thailand views Mathematics as an application of knowledge, skills and scientific process for problem solving, as a way of life and to further education, and aims at developing students who are systematic and constructive thinkers (Thailand. Ministry of Education, 2008). The Basic Education Core Curriculum implemented in 2008 stresses mathematical thinking. The Mathematical skills and processes stipulated in the curriculum include problem solving, reasoning, communication, representation, connecting, and creativity (Thailand. Ministry of Education, 2001, 2008). The curriculum covers six content strands (*Number and Operations*, *Measurement*, *Geometry*, *Algebra*, *Data Analysis and Probability,* and *Mathematical Skills and Processes-Problem Solving*). Every content strand focuses on the importance of students connecting various bodies of mathematical knowledge and connecting mathematical knowledge with other disciplines, and attaining the ability for creative thinking.

The learning strand on *Mathematical Skills and Processes- Problem Solving* is specifically targeted on developing students' mathematical skills and processes that are related to problem solving (Dechsri, 2012). This content strand addresses the diverse methods that can be used to solve problems, which includes communication capacity, thinking capacity, problem solving capacity, capacity for applying life skills and capacity for technical application. Among these five competencies, the first three (communication capacity, thinking capacity, problem solving capacity) are directly related and relevant to mathematics learning In the communication capacity, the emphasis is one the ability to distinguish and select information through reasoning, while the thinking capacity focuses on analytical, synthetic, constructive, critical and systematic thinking. The problem solving capacity stresses on problem solving skills and decision making (Thailand. Ministry of Education, 2008). Based on the released report entitled 'Analysis of Upper Secondary Mathematics Standards from Seven Asia-Pacific Economic Cooperation (APEC) Economies', it was identified that one third of the standards found in Thailand Mathematics Curriculum is

focused on the two levels of cognitive demand- conjecture/generalise/prove and solve non-routine problems (Robertson-Kraft & Porter, 2011) which are associated to HOT.

In comparison, only Thailand dedicates one learning area for the mathematical skills and processes related to problem solving, while the other three countries incorporate the skills and processes related to problem solving in all the content strands.

To summarise, the mathematics curriculum of these four countries focuses on problem solving skills to varying degrees and implicitly state the cognitive processes related to HOT such as reasoning, making decisions or interpreting, in addition to LOT.

Research Questions 2 and 3

The percentage of students at each benchmark and students' percentile scores were computed for each of the three cognitive domains for the four countries. The percentage of students at each benchmark informs the proportion of students who answered the items assessing the three cognitive domains, which are relevant to determining the proportion of students who were able to answer LOT and HOT items. The analyses on the nth percentile scores provides information on the highest score obtained by the bottom n% of the students (for $5 \leq n < 50$) and the lowest score obtained by the top n% of the students (for $50 < n \leq 95$). As such, these two analyses are vital in investigating students' performance on items that assess HOT and LOT.

The cumulative percentage of students at each benchmark 'at 400 or above' for low 'at 475 or above' for intermediate, 'at 550 or above' for high and 'at 625 or above for advanced) was transformed to obtain the percentage of students at each benchmark. These benchmarks were then relabelled as low (obtaining score points 400-474), intermediate (475-549 points), high (550-624) and advanced (at least 625 points). Another category was also added to get the percentage of students below the 'low category' and was labelled as 'weak'(less than 400 points). Table 5 shows the percentage of students at each international benchmark for all the four countries for the three cognitive domains.

The percentages for each domain were compared relatively across the three domains. At this point, it is important to highlight that the weak category represents category of students who obtained below 400 score points and as such, need to be carefully interpreted. This is because smaller values are desired as it indicates a smaller group obtaining score

points of below 400. This interpretation is only true for the weak category and not for the other categories.

Table 5
Percentage of Advanced, High, Intermediate, Low and Weak Students

Cognitive Domains	Country	Student % below Low Benchmark **(Weak)** < 400 points	Student % at Low Benchmark **(Low)** 400-474 points	Student % at Intermediate Benchmark **(Intermediate)** 475-549 points	Student % at High Benchmark **(High)** 550-624 points	Student % at Advanced Benchmark **(Advanced)** ≥ 625 points	Σ
	Indonesia	59.53	25.19	11.79	2.98	0.51	100.00
	Malaysia	33.94	27.46	23.71	12.29	2.60	100.00
Knowing	Singapore	0.92	5.40	13.86	29.29	50.53	100.00
	Thailand	41.36	31.96	18.69	5.94	2.05	100.00
	Indonesia	57.17	28.13	12.19	2.23	0.28	100.00
	Malaysia	35.74	28.38	23.49	10.93	1.46	100.00
Applying	Singapore	1.30	6.16	14.81	29.25	48.48	100.00
	Thailand	38.14	33.66	20.50	6.19	1.51	100.00
	Indonesia	55.21	29.60	12.60	2.37	0.22	100.00
	Malaysia	40.78	27.66	20.56	9.26	1.74	100.00
Reasoning	Singapore	2.88	7.81	16.08	27.44	45.79	100.00
	Thailand	38.20	31.42	21.11	7.43	1.84	100.00

Based on these values, a line and bar graph were drawn for each country to illustrate the percentage of students at each benchmark for easy comparisons between LOT and HOT items. Figures 1 to 5 display the graphs that were plotted, where the y-axis represents the percentage of students and the x-axis represents the three cognitive domains for each student category.
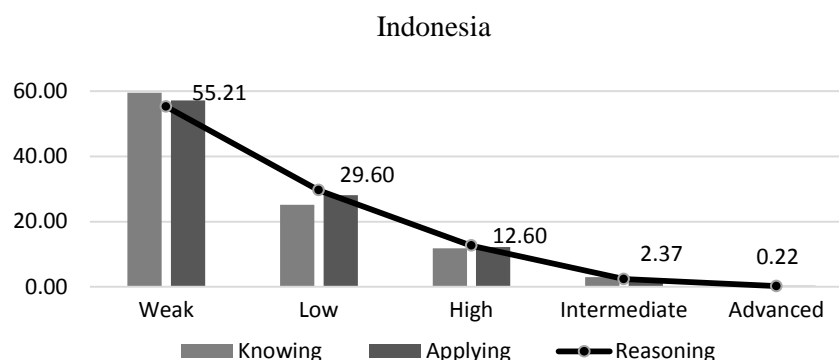
Figure 1. Percentage of Indonesian Students at each Benchmark

In Figure 1, it can be deduced that for the weak category, there is a higher percentage of students who obtained score less than 400 in the *knowing* and *applying* domains of LOT than in the *reasoning* domain of HOT. This suggests that there is a higher proportion of students in the weak category who were able to answer HOT items than LOT items. In all other student categories of low, high, intermediate and advanced, there is a higher percentage of students who were able to answer HOT items (*reasoning* domain) than in the LOT items (*knowing* and *applying* domains). The conclusion is that in Indonesia, there is generally a higher student proportion who were able to answer HOT items.
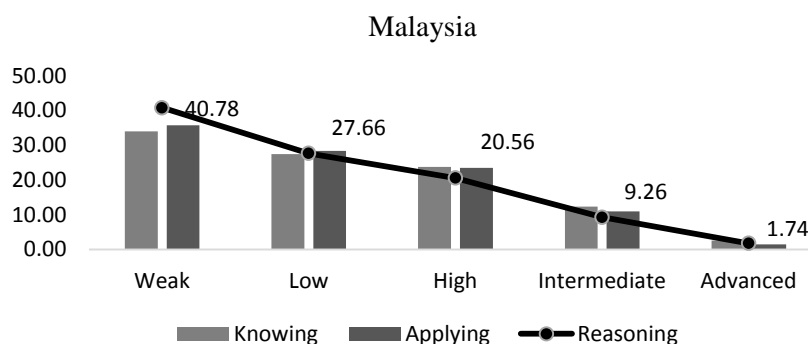

Figure 2. Percentage of Malaysian Students at each Benchmark

From Figure 2, there are more students in the weak category who obtained below 400 points in the *reasoning* domain than in the *knowing* and *applying* domains, indicating that more Malaysian students at the weak category were more able to answer LOT than HOT items. For all other categories of low, intermediate, high and advanced, there is a higher proportion of students in the *knowing* and *applying* domains than in the *reasoning* domain. This again indicates that a higher proportion of Malaysian students were able to answer LOT items than HOT items. The conclusion is that across all the categories of ability, a higher proportion of Malaysian students were able to answer LOT when compared to HOT items.
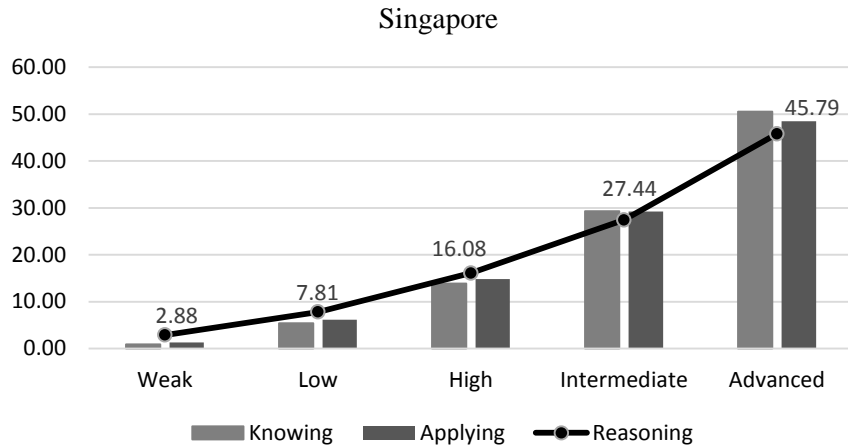
Figure 3. Percentage of Singaporean Students at each Benchmark

For Singapore students, two clear patterns seem to emerge as seen in Figure 3. For the categories of low and high, a bigger percentage of students is located at the *reasoning* domain than at the *knowing* and *applying* domains. This observation suggests that more students at the low and high categories were able to answer HOT items than LOT items. Conversely, for the categories of weak, intermediate and advanced, a higher proportion of students are concentrated at the *knowing* and *applying* domains than at the *reasoning* domain, with minimal percentage difference for the intermediate category. This indicates that a higher student proportion at the intermediate and advanced categories were able to answer LOT items than HOT items. Similarly for the weak category, there is a higher percentage of students who scored below 400 points at the *reasoning* domain than at the *knowing* and *applying* domains, which also suggests that there is a higher proportion of students in the weak category who were able to answer LOT items than HOT items. The conclusion is a higher proportion of Singaporean students at the low and high categories were able to answer HOT items, while a higher proportion of students at the weak, intermediate and advanced categories were able to answer LOT items.
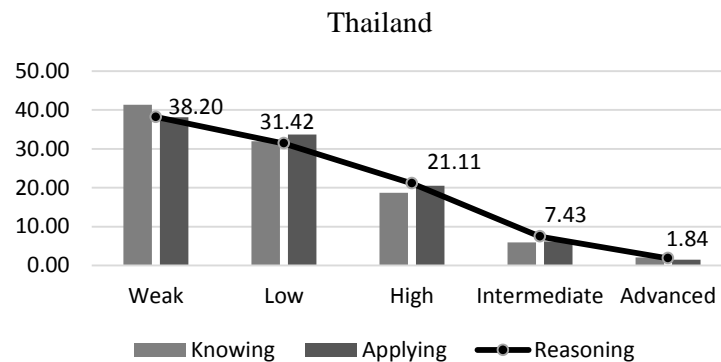
Figure 4. Percentage of Thai Students at each Benchmark

As can be seen in Figure 4, there are more Thai students in the weak, high, intermediate and advanced categories at the *reasoning* domain than at the *knowing* and *applying* domains. For the weak category, there is a smaller percentage of students at the *reasoning* domain than at *knowing* and *applying* domains. These observations indicate that students at the categories of weak, high, intermediate and advanced were inclined towards answering HOT items than LOT items. Only for the low category, the student proportion at the *knowing* domain is slightly lower than at the *applying* domain and almost similar when compared to the *reasoning* domain. This may also suggest that students at the low category were also inclined towards answering HOT items. The conclusion is generally a higher proportion of Thai student were able to answer HOT items than LOT items.

In order to further investigate students' performance, percentile (5th, 10th, 25th, 50th, 75th, 90th and 95th) analyses were carried out. The analyses obtained were used to plot a bar and line graph as shown in Figures 6-10.

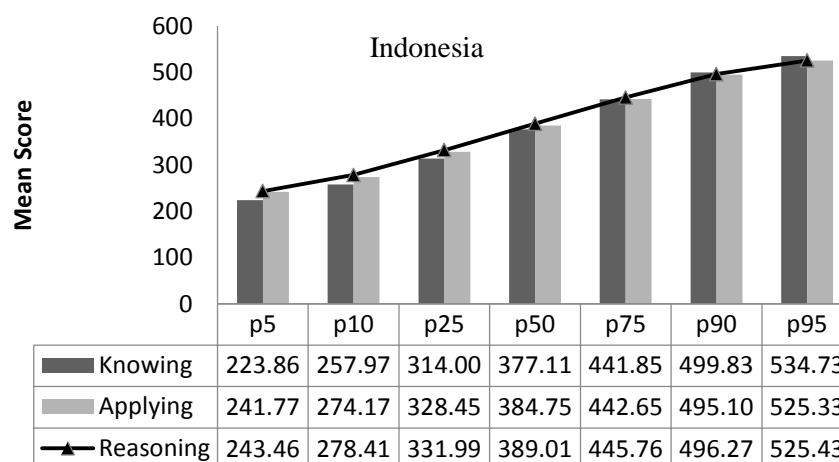|	p5	p10	p25	p50	p75	p90	p95
Knowing	223.86	257.97	314.00	377.11	441.85	499.83	534.73
Applying	241.77	274.17	328.45	384.75	442.65	495.10	525.33
Reasoning	243.46	278.41	331.99	389.01	445.76	496.27	525.43

Figure 5. Percentile Achievement for Indonesian Students for Cognitive Domains

As displayed in Figure 5, from the 5th percentile to the 75th percentile, the highest score obtained by the bottom 5%, 10 %, 25%, and 50%, and the lowest mark obtained by the top 25% is higher for items at the *reasoning* domain compared to *knowing* and *applying* domains, suggesting students' better performance on HOT items. Even though the 90th and 95th percentiles show that the lowest scores for the top 10% and top 5% respectively are slightly lower for items at the *reasoning* domain compared to *knowing* and *applying* domains, the score difference between the domains is minimal. Therefore, indicating that these students answered equally well on HOT and LOT items. The conclusion is that Indonesian students scored better on HOT items, except for the high achievers whose performance were similar on LOT and HOT items.
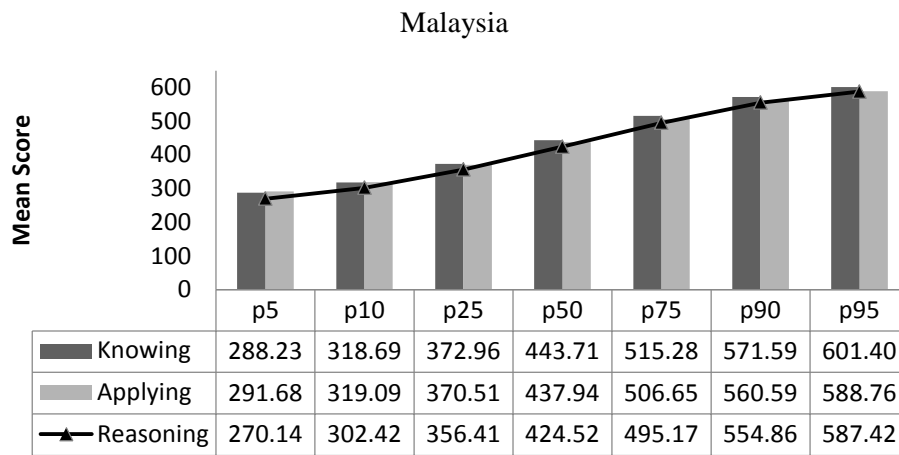

Malaysia

	p5	p10	p25	p50	p75	p90	p95
Knowing	288.23	318.69	372.96	443.71	515.28	571.59	601.40
Applying	291.68	319.09	370.51	437.94	506.65	560.59	588.76
Reasoning	270.14	302.42	356.41	424.52	495.17	554.86	587.42

Figure 6. Percentile Achievement for Malaysian Students for Cognitive Domains

From Figure 6, unlike Indonesia, Malaysian students at every percentile performed better on items at the *knowing* and *applying* domains compared to *reasoning* domain. This suggests that Malaysian students scored higher on items assessing LOT than HOT.
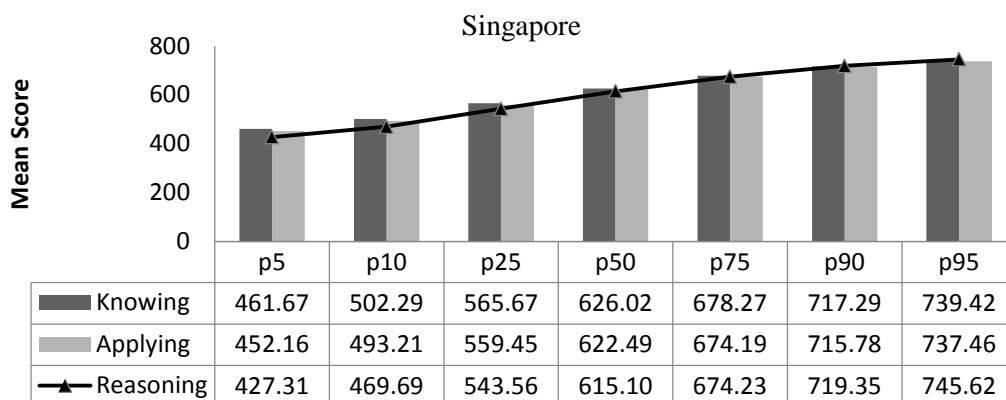

Singapore

	p5	p10	p25	p50	p75	p90	p95
Knowing	461.67	502.29	565.67	626.02	678.27	717.29	739.42
Applying	452.16	493.21	559.45	622.49	674.19	715.78	737.46
Reasoning	427.31	469.69	543.56	615.10	674.23	719.35	745.62

Figure 7. Percentile Achievement for Singapore Students for Cognitive Domains

Figure 7 shows that the highest score for the bottom 5%, 10%, 25% (5th , 10th and 25th percentiles) are higher for items at the *knowing* and *applying* domains than for items at the *reasoning* domain. This indicates that the students scored better on LOT items. Noticeably, the lowest scores for the top 10% and 5% of the students (90th and 95th) are higher for items at the *reasoning* domain than the *knowing* and *applying* domains, which suggests that the top achievers obtained higher score on HOT items than LOT items. The conclusion is except for the low achievers, Singapore students scored better on HOT items than on LOT items
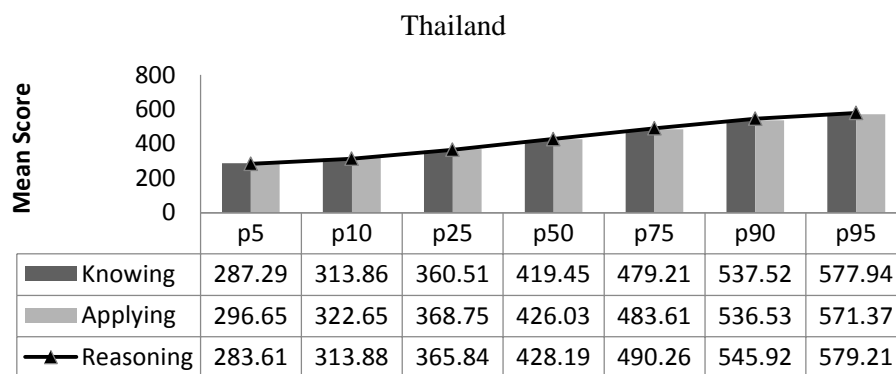
Thailand

	p5	p10	p25	p50	p75	p90	p95
Knowing	287.29	313.86	360.51	419.45	479.21	537.52	577.94
Applying	296.65	322.65	368.75	426.03	483.61	536.53	571.37
Reasoning	283.61	313.88	365.84	428.19	490.26	545.92	579.21

Figure 8. Percentile Achievement for Thai Students for Cognitive Domains

As exhibited in Figure 8, the highest score for the bottom 5% (5th percentile) until the lowest score for the top 5% (95th percentile) is higher for items that assess *reasoning* than items that assess *knowing* and *applying* skills. This observation suggests that the general Thai student population obtained higher score on HOT items than LOT items, which brings to the conclusion that Thai students consistently obtained higher score on HOT items than on LOT items.

Conclusion and Discussion

The National Council of Teachers of Mathematics (2000) stipulates five major mathematical processes of acquiring and applying the mathematical knowledge. They are problem solving, reasoning and proof, communication, connection and representation. In all these four countries, it is apparent that these five cognitive processes are reflected in the curriculum, even though not all these processes are clearly stipulated in the curriculum. By inspecting the intended curriculum as stipulated in the curriculum of these four countries, it is apparent that Indonesia and Singapore stresses connecting the mathematics learnt in the classroom with real life situations, while Thailand dedicates one learning area focussed primarily on problem solving processes and skills, which includes communication capacity,

thinking capacity and problem solving capacity. The mathematics curriculum in Malaysia emphasises problem solving just like in the curriculum adopted by the other three countries.

By studying the results from the benchmark and percentile analyses, it can be concluded that a higher proportion of Indonesian students were able to answer HOT than LOT items, and the percentile scores were also higher for HOT items than LOT items. The general Malaysian student population consistently did better on LOT items than HOT items as indicated by the high percentage of students and the low percentile score for HOT items. Except for the low achievers (bottom 25%), Singapore students generally performed better on HOT than LOT items. However, it is noted that while the top achievers recorded a higher percentile score on HOT items, the proportion of students at the advanced category was smaller when compared to the proportion of students who were able to answer LOT items. For the Thai students, they generally performed better on HOT items as testified by the higher student percentage and higher percentile score for HOT when compared to LOT items.

By studying these findings within the context of the curriculum adopted in these four countries, Indonesia, Singapore and Thailand performed well on HOT items than on LOT items, at varying degrees. On the other hand, Malaysian students performed consistently better on LOT items than HOT items.

This raises the question why, especially when the mathematics curriculum implemented in all these four countries emphasise problem solving. Do the findings indicate that emphasising on problem solving alone is inadequate? Could the answer lie in the emphasis of relating the mathematics learnt in the classroom to connecting it to real life situations through non-routine questions or dedicating one learning area to skills related to problem solving or beyond that, the teaching styles that focus on real-life situations and create a problem solving culture in the classroom? By connecting mathematics to real life, does it enhance students' HOT skills as they are consistently exposed to situations where they are required to apply the mathematics learnt, to not only the classroom or 'text book' problems, but in real life situations? Does using open-ended mathematics questions 'stretch' students' mind?

The common denominator appears to point to the emphases on connecting the mathematics to real life situations, and strategize the instructional activities during mathematics lessons to facilitate students' problem solving through the mathematical skills

and processes. The author is in the opinion that connecting mathematics to real life situations opens up a getaway of exposing students to a host of innumerable non-routine questions and divergent opportunities to use their 'extended mind' in productively solving unfamiliar tasks using past experiences that were not previously associated. The novel situations set in the non-routine questions and the need to connect the learnt mathematics to real life situations compel students to switch from rehearsed algorithms into thinking 'out of the box'. As they manipulate information to explore a plethora of correct and incorrect algorithms, they are directly and intensively engaged in reasoning out their answers and making decisions to accept or disapprove those answers. The mathematical processes and skills acquired during this process of reasoning 'stretch' the use of mind. Therefore, they are crucial in developing students' higher order thinking skills.

This findings of this study ignite a possibility of emphasising 'making real-life connections' through non-routine questions, and infusing the mathematics lessons with the mathematical skills and processes associated to problem solving as a pedagogical practice to encourage students to think and reason out and thus, enhance HOT skills. However, more studies need to be conducted in this direction before making any conclusions.

References

Chin, T.Y. et al. (2012). Singapore. In Mullis, I.V.S. et al. (Eds.), *TIMSS 2011 Encyclopedia Vol 2: Education policy and curriculum in Mathematics and Science* (pp. 801-815). Chesnut Hill, MA: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.

Dechsri, P. (2012). Thailand. In Mullis, I.V.S. et al., (Eds.), *TIMSS 2011 Encyclopedia Vol 2: Education policy and curriculum in Mathematics and Science* (pp. 907-921). Chesnut Hill, MA: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.

Maier, N.R.F. (1933). An aspect of human reasoning. *British Journal of Psychology, 24,* 144-155

Malaysia. Ministry of Education. (2004a). *Integrated curriculum for secondary schools: Curriculum specifications for Mathematics Form 2*. Kuala Lumpur, Malaysia: Curriculum Development Centre.

Malaysia. Ministry of Education. (2004b). *Integrated curriculum for secondary schools: Syllabus mathematics*. Kuala Lumpur, Malaysia: Curriculum Development Centre.

Muhammad Zaini Mohd Zaini & Dewani Goloi (2012). Malaysia. In Mullis, I.V.S. et al. (Eds.), *TIMSS 2011 Encyclopedia Vol 2: Education policy and curriculum in Mathematics and Science* (pp. 569-584). Chesnut Hill, MA: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.Mullis, I.V.S., Martin, O.M., Ruddock, G. J., Sullivan, C.Y. & Preuschoff, C. (2009). *TIMSS 2011 Assessment Frameworks*. Chesnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Mullis, I.V.S., Martin, M.O., Foy, P. & Arora, A. (2012). *TIMSS 2011 International results in Mathematics*. Chesnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Mullis, I. V. S., Martin, M. O., Minnich, C. A., Stanco, G. M., Arora, A. A., Centurino, V. A. S. & Castle, C.E. (2012a). *TIMSS 2011 Encyclopedia: Education policy and curriculum in Mathematics and Science* (Vol. 1). Chesnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Mullis, I. V. S., Martin, M. O., Minnich, C. A., Stanco, G. M., Arora, A. A., Centurino, V. A. S. & Castle, C. E. (2012b). *TIMSS 2011 Encyclopedia: education policy and curriculum in Mathematics and Science* (Vol. 2). Chesnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

National Council of Teachers of Mathematics (2000). *Principles and Standards for School Mathematics.* Retrieved from http://www.nctm.org/uploadedFiles/Math_Standards/12752_exec_pssm.pdf

Newman, F.M. (1990). Higher order thinking in teaching social studies: A rational for the assessment of classroom thoughtfulness, *Journal of Curriculum Studies, 22*, 41-56.

Quellmalz, E. S. (1985). *Needed: Better methods for testing higher-order thinking skills*. Retrieved from http://www.ascd.org/ASCD/pdf/journals/ed_lead/el_198510_quellmalz.pdf

Rahmawati, B. (2012). Indonesia. In Mullis, I.V.S. et al. (Eds.), *TIMSS 2011 Encyclopedia Vol 1: Education policy and curriculum in Mathematics and Science* (pp. 397-405). Chesnut Hill, MA: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.

Rajendran, N.S. (2010). *Teaching & acquiring higher-order thinking skills.* Perak, Malaysia: Penerbitan Universiti Pendidikan Sultan Idris.

Reba'i, M. (2013). *Apa, mengapa dan bagaimana Pendidikan Matematika Realistik Indonesia (PMRI).* Retrieved from http://www.academia.edu/5978707/APA_MENGAPA_DAN_BAGAIMANA_PMRI

Robertson-Kraft, C. & Porter, A. (2011). *Analysis of Upper Secondary Mathematics Standards from Seven Asia-Pacific Economic Cooperation (APEC) Economies.* Retrieved from http://www.cgp.upenn.edu/ope/documents/APEC_Math_Standards.pdf

Sembiring, R. K., (2010). Pendidikan Matematika Realistik Indonesia (PMRI): Perkembangan dan Tantangannya. *IndoMS. J.M.E, 1* (1), 11-16

Sembiring, R.K., Hadi, S., & Dolk, M. (2008). Reforming mathematics learning in Indonesian classrooms through RME. *ZDM Mathematics Education, 40,* 927–939

Singapore. Ministry of Education. (2006). *Secondary mathematics syllabus. Singapore: Curriculum planning and development division.* Retrieved from www.moe.gov.sg/education/syllabuses/sciences/files/maths-secondary.pdf

Thailand. Ministry of Education. (2008). *The basic education core curriculum B.E. 2551 (A.D. 2008).* Retrieved from http://www.act.ac.th/document/1741.pdf

Thailand. Ministry of Education. (2001). Curriculum of basic education 2001. Retrieved from http://nrru.ac.th/be/data/cur.pdf.

Thailand. Ministry of Education. (2008). *Core curriculum of basic education 2001.* Retrieved from http://www.khaotolon.com/images/sub_1218591375/intro_mid_2551.pdf

Thompson, T. (2 011). *An analysis of higher-order thinking on Algebra I end-of course tests.* Retrieved from   http://www.cimt.plymouth.ac.uk/journal/thompson.pdf