Assessing ASEAN Students' Competencies and Readiness in Learning Mathematics Through MaRWA Diagnostic Test

Wisnuningtyas Wirani, SEAMEO QITEP in Mathematics, Indonesia
Ummy Salmah, SEAMEO QITEP in Mathematics, Indonesia
Bungkus Dias Prasetyo, SEAMEO QITEP in Mathematics, Indonesia
Uki Rahmawati, SEAMEO QITEP in Mathematics, Indonesia

The Asian Conference on Education & International Development 2023
Official Conference Proceedings

Abstract
MaRWA is a Southeast Asia regional mathematics assessment based on PISA and TIMSS, initiated by SEAMEO QITEP in Mathematics. This paper aims to determine students' type of correct and incorrect answers on the assessment. A total of 882 students in grades 5, 8, and 10 from 26 schools in regional wide were involved in this study. For each level of education, there were 30 mathematics problems, which were classified into 20 multiple choice questions and 10 essay questions. These items were generally made to determine students' competencies and readiness in learning school mathematics, specifically to measure their order of thinking level. There were four types of correct answer and six types of incorrect answers. The quantitative data is the students' scores, while the qualitative data is coded students' answers. It was found that students' average scores for the primary to senior high school levels were in a low category, and most of the participants were struggling to solve HOTS problems. Moreover, most students in all levels left problems in blank. We foresee this study can contribute to bridge the students' HOTS by improving the practice of teaching and learning throughout Southeast Asia.

Keywords: Mathematics Assessment, Types of Error, Higher Order Thinking Skill, Mathematics Teaching and Learning, ASEAN

Introduction

Introduction to SEAMEO QITEP in Mathematics

The Southeast Asian Ministers of Education Organization (SEAMEO) is a regional intergovernmental organisation established in 1965. Its objective is to nurture cooperation among Southeast Asian countries governments in three areas named education, science, and culture. For education, there are seven priority areas: achieving universal early childhood care and education; addressing barriers to inclusion; promoting resilience in the face of emergencies; promoting technical and vocational education and training; revitalising teacher education; harmonising higher education and research; and adopting a 21st century curriculum.

The executive arm of the Council and the centre of SEAMEO organisations is SEAMEO Secretariat which located in Bangkok, Thailand. Among 26 SEAMEO regional organisations spread in Southeast Asian countries, three regional centres are categorised in the education cluster and focused on improving teachers and education personnel quality. One of which is SEAMEO Regional Centre for Quality Improvement of Teachers and Education Personnel (QITEP) in Mathematics (SEAQiM) which is located in Yogyakarta, Indonesia. This institution has major goals such as catering for the needs to improve the quality of mathematics teachers and education personnel in Southeast Asia; promoting sustainable teacher professional development in the area of mathematics education; establishing extensive networks, information exchange, and best practice sharing in mathematics education among SEAMEO member countries; and providing intellectual forums on mathematics education innovation (SEAQiM, 2012).

SEAQiM runs many programmes to improve mathematics teachers' quality throughout Southeast Asia. It has a flagship programme named regular course with different themes such as Differentiated Instruction, Lesson Study, Realistic Mathematics Education, STEM Education, Joyful Learning, Teacher-Made Teaching Aids, and Clinical Supervision. The regular course usually holds three to four times in a year.

Other than course, SEAQiM also develop an assessment for mathematics learning, especially in determining students' higher order thinking skills (HOTS) in learning mathematics. This programme is called Mathematics Regional Wide Assessment (MaRWA) and will be described in the next section.

MaRWA

MaRWA is a Southeast Asia regional assessment which aims at determining students' readiness in learning mathematics. It is a diagnostic test for students across Southeast Asia in grade 5 (primary level), grade 8 (junior high level), and grade 10 (senior high level). This programme was established in 2012, considering the importance of regional benchmarking.

Why developing MaRWA? Why do not we just believe the result of other international assessment such as The Programme for International Student Assessment (PISA) or Trends in International Mathematics and Science Study (TIMSS)? Both PISA and TIMSS are international assessments. In PISA, the test items are different to what students usually have in school as the test is not about remembering what students have learned but school, but how students can use their knowledge for problem solving (Schleicher, 2019). There are critics

addressed to PISA (Hopmann, 2008; Eivers, 2010; Sjøberg, 2015; Forestier & Adamson, 2017; Zhao, 2020). Some critics are regarding its implementation, statistical analysis, interpretations, and immense influence on a country's education. Another critique is on the use of context for the problems which is incompatible with the culture of the country-test-taker as assessing mathematics achievement is related to a country's educational activities (Lamichhane, 2018). Thus, to ascertain of more valid and constructive results, a country should develop its own localised assessment, in which its contexts and characteristics is in accordance with a country's culture (White, 2017; White, 2018).

Developed items for MaRWA used PISA and TIMSS as the references. However, one of important points is that, MaRWA test items should not duplicate test items on PISA and TIMSS as it is an ASEAN localised assessment. Moreover, this is a diagnostic test, not a test to rank participants as its result will be used to detect students' strengths and weaknesses in learning mathematics. Through the result, MaRWA is expected to help mathematics teachers in recognising the problems of students learning and provide inputs for teacher education providers to conduct relevant teacher professional development for mathematics teachers in Southeast Asia.

Developing MaRWA Items

In 2012, numbers of experts in mathematics assessment from different countries, for instance Australia, Japan, Malaysian, and Indonesia, developed MaRWA items for grade 5 and grade 8 in the form of multiple choice, explanation, and short answer. As previously mentioned, MaRWA items were referring to PISA, in which to assess students' HOTS. These cover aspects of HOTS such as students' ability to select relevant information to solve problems, to use different strategies, and to utilise critical thinking and metacognition (OECD, 2021).

After developing these items, a pilot was conducted in 2013 in Indonesia, Vietnam, Philippines, and Cambodia. It was resulted in items which were readable and had a good discriminant factor. Then, an extended try-out was employed in 2014 at several schools in Cambodia, Indonesia, Thailand, Timor Leste, and Vietnam. More items were developed in 2016, and on that year the level was added, grade 10 senior high. These items were tested in schools in Indonesia through paper test and online test. From then on, SEAQiM has committed to implement the assessment online.

Error Categories

Newman (1977, as cited in Clements, 1980) created five error categories: encoding, process skills, transformation, comprehension, and reading. Problems in MaRWA involves contexts. Therefore, we used comparation of Newman's error categories and Blum and Leiss' modelling process (Wijaya, van den Heuvel-Panhuizen, Doorman, & Robitzsch, 2014). It turned out that the error categories and the coding is in line with the modelling process and mathematisation stages of PISA. To determine students' error, we used error categories of Wijaya et al. (2014) which classified the errors into comprehension, transformation, mathematical processing, encoding, and unknown. Compared to Newman's and Wijaya's, who constructed five error categories, we also added an error type in which students left the problem in blank, did not answer the problem, or gave irrelevant answer. This will be written in details in Methodology section.

Based on the abovementioned sections, this study aims at examining ASEAN students' competencies in working with mathematics problems of different levels thinking skills through MaRWA. Therefore, in this study we expect to determine what is the common type of error and correct students' answers, especially on HOTS items, in MaRWA.

Methodology

Method

This study employed case study research design. According to Creswell (2007), case study is a qualitative approach to determine a specific case for periods of time. Moreover, this research design is appropriate for exploring a phenomenon in a real-life context (Yin, 2003). Mixed method is commonly used in case study research design. Therefore, in this study we analysed the collected data quantitatively and qualitatively. This paper was focused on identifying the most common type of errors and correct of students' answers during MaRWA 2020.

Participants of MaRWA 2020

A total of 882 students in grades 5, 8, and 10 from 26 schools in regional wide were involved in this study. These 26 schools were located in three countries: Indonesia, Malaysia, and Philippines. In details, there were 204 students grade 5 from 10 schools in Indonesia, Malaysia, and Philippines; 169 grade 8 students from nine schools in Indonesia; and 509 grade 10 students from seven schools in Indonesia.

Test Items

For each level of education, there were 30 mathematics problems, which were classified into 20 multiple choice questions and 10 essay questions. These items were generally developed to determine students' competencies and readiness in learning school mathematics, specifically to measure their order of thinking level. Three levels of thinking: higher order thinking skill (HOTS), middle order thinking skill (MOTS), and lower order thinking skill (LOTS); were embedded in the set of problems. Items of MaRWA covered strands such numbers, geometry and measurement, algebra, and statistics. In details, the strands and levels are presented in Table 1.

Table 1: Strands of MaRWA Items for Each Level

Level	Strands
Primary	Numbers and operation
	Quantity and measurement
	Shapes, figures, and solids
	Data handling and statistics
Junior High	Numbers
	Algebra
	Geometry
	Statistics and probability
Senior High	Algebra
	Geometry
	Trigonometry
	Statistics and probability

These items were developed and revised in 2012 and 2016. As MaRWA has been employed for more than 10 years, these mathematics problems have been tested to more than 200 students in Southeast Asia region on each level.

Test Implementation

The test of MaRWA 2020 was carried out on June to July 2020, it was during COVID-19 pandemic. Moreover, as the budget of MaRWA implementation was limited, we digitalised the developed test items and uploaded these into a platform named Quia. The platform is suitable for test items in the form of multiple choice and essay. It enables students in writing their arguments in details.

To take the test, the mathematics teacher registered the school by sending an email to marwa@qitepinmath.org to set the test date. Before the test, SEAQiM ensured that the internet connection in the school was stable through confirmation of the mathematics teacher. Once the date set, SEAQiM shared the test link, it was different links for each school, and students will be given 90 minutes to work on the test. During the test, students had to fill in their names.

Analysing Test Items

The data in this study is students' answers. As mentioned before, the test items of MaRWA included multiple choice and essay questions. There were two ways in analysing students' answer. For multiple choice answers, we utilised statistical analysis such as finding out the maximum, the minimum, and the average score by using Excel. As for essay questions, we distinguish between correct and incorrect answer. We categorised answers on essay questions based on classification of correct and incorrect (error) answers as displayed in Table 2 and Table 3. To add, we also utilised statistical analysis to determine percentage of correct and incorrect answers types on essay questions.

Table 2: Correct Types and Codes for Analysing MaRWA Results

Type	Coding
Student only writes the final answers	B0
Student writes incorrect strategy/procedure	B1
Student writes mysterious or undefined strategy/procedure	B2
Student writes correct strategy/procedure	B3

Table 3: Error Types and Coding Scheme for Analysing MaRWA Results

Error Type	Sub-type	Explanation	Coding
Comprehension	Misunderstanding the instruction	Student incorrectly interprets what s(he) is asked to do	E1
	Misunderstanding a keyword	Student misunderstands a keyword, which is usually a mathematical term	

	Error in selecting information	Students is unable to distinguish between relevant and irrelevant information (e.g using all information provided in a task or neglecting relevant information) or is unable to gather required information which is not provided in the task.	
Transformation	Procedural tendency	Student tends to directly use a mathematical procedure such as formula or algorithm without analysing whether or not it is needed.	E2
	Taking too much account of the context	Student's answer only refers to the context or real-world situation without taking the perspective of the mathematics.	
	Wrong mathematical operation/concept	Student uses mathematical procedures or concepts which are not relevant to the tasks.	
	Treating a graph as a picture	Student treats a graph as a literal picture of a situation. (S)he interprets and focuses on the shape of the graph, instead of on the properties of the graph.	
Mathematical Processing	Algebraic error	Error in solving algebraic expression or function	E3
	Arithmetical error	Error in calculation	
	Error in mathematical interpretation of data representation (e.g chart, graph, etc.)	Student mistakenly focuses on a single point rather than on an interval.	
		Student does not use the slope of the graph but only focuses on the vertical distance.	
	Measurement error	Student cannot convert between standard units (e.g. from m/minute to km/h) or from a non-standard unit to a standard unit (e.g. from step/minute to m/minute) and rounding errors.	

	Improper use of scale	Student cannot select and use the scale properly.	
	Unfinished answer	Student uses a correct procedure, but (s)he does not finish it.	
		Student is not able to answer the final answer that is caused by the previous error in the problem-solving.	
Encoding		Student is unable to correctly interpret and validate the mathematical solution in terms of the real-world problem. This error is reflected by an impossible or not realistic answer	E4
Unknown		Type of error could not be identified due to limited information from student's work.	E5
No Answer/Invalid Answer		Students did not answer the problem or gave irrelevant answer	E6

To determine the category for students' average score, we used Table 4 to describe.

Table 4: Categories for Students' Average Score

Category	Average Score
Very high	86 – 100
High	76 – 85.99
Moderate	60 – 75.99
Low	55 – 59.99
Very low	< 54

Results and Discussion

In this section, we will answer the research question "What is the common type of error and correct students' answers, especially on HOTS items, in MaRWA?" based on findings on each level. To add, we will also present examples of students' answers based on the level of students thinking skills in three categories: HOTS, MOTS, and LOTS.

Primary Level

From 30 items, there were 8, 14, and 8 questions for HOTS, MOTS, and LOTS, respectively. In other words, the number for MOTS items is the greatest of all. According to the analysis, it was found that the students' average score is very low. Furthermore, compared to other levels, primary school average score is the highest among all. The statistical result for students in primary school is displayed in Table 5.

Table 5: Statistical Data on Primary Level

Statistical Data	Score (maximum of 100)
Maximum	83.33
Minimum	0
Average	26.54

To have a deeper comprehension on students' answer for essay questions which contains three levels of thinking skills, we ran statistical analysis. Figure 1 shows a pie chart on the percentage of correct answer types. We found that most students who answered correctly on essay questions were able to explain in an appropriate strategies or procedures (B3). Based on the findings, less than 2% of primary students were able to write a correct final answer but failed to provide appropriate strategies or procedures (B1).

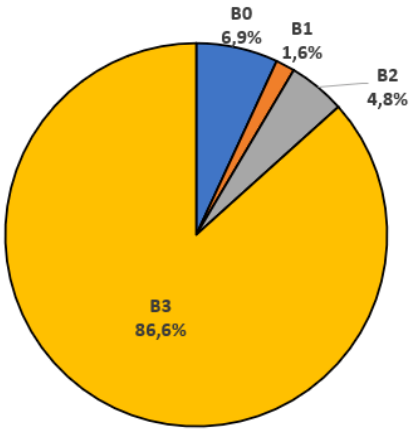

Figure 1: The percentage for each correct answer type on primary level

Not only statistical analysis for correct answer, we also ran for incorrect answer on essay questions as showed in Figure 2. According to findings, most primary students were unable to provide answer or they provided invalid answer (E6). Another error type that usually occurred is that students' strategy could not be identified due to limited explanation (E5).

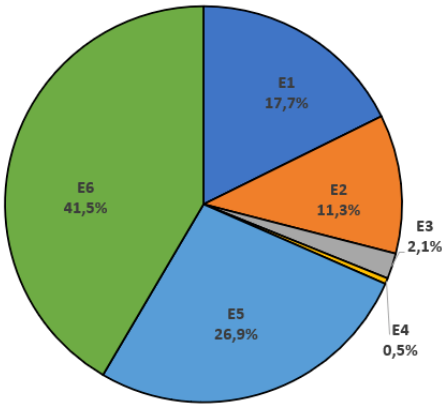

Figure 2: The percentage for each error type on primary level

According to statistical analysis, the correct and error type for HOTS items is in line with findings the percentage for correct and error type on essay questions. For HOTS items on essay questions, most students were fall into error type E6 (43%) and correct type B3 (74.4%).

HOTS Items for Primary Level

Figure 3 illustrates the percentage of students' answer for HOTS items. More than 70% of students on primary level were unable to answer correctly for all HOTS items. These top three items: question 3, question 17, and question 29, were HOTS items with the highest unsolved percentage of all.

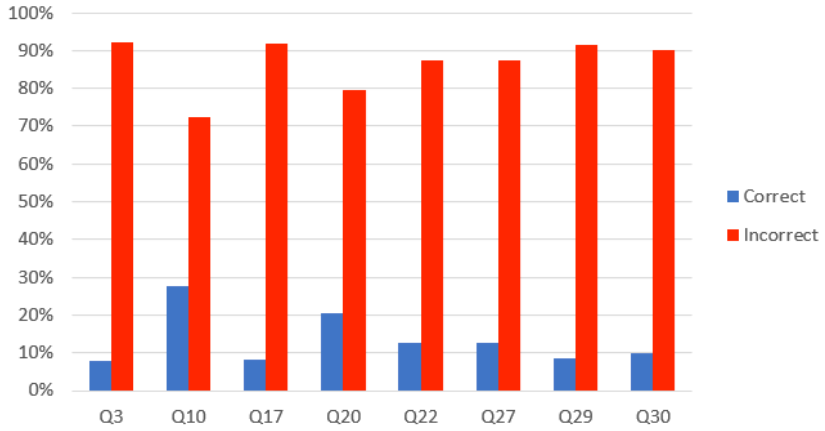

Figure 3: The percentage of primary students' answer for HOTS items

To illustrate, we provide question 29 and samples of students' answer in Figure 4.

29. A container filled up with water is tilted as shown in the picture. How much water is thrown out from the container?
(1 point)

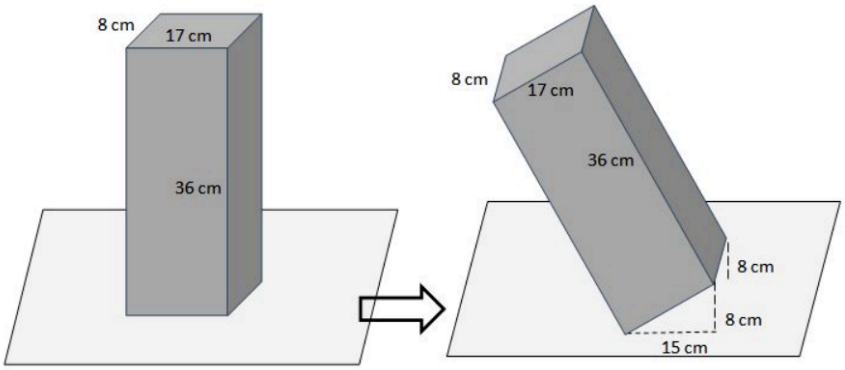

Figure 4: A sample of HOTS item for primary level, question 29

Samples of students' error:

- 18×17×36×8×8×15=4.700.160 (E1)
- 8×17×36 = 4896
  8×15×36 = 4320
  4896 – 4320 =576 (E1)

These two samples are illustrating the error type of comprehending the problem (E1). The problem is asking for amount of water that is thrown out of the container when it is tilted. However, the first sample provides us an information that the student was unable to use information to solve the problem. Rather than determining the volume of a prism then subtracting it with the amount of thrown water, this student multiplied all numbers on the question. The student has failed to comprehend in determining the volume of a prism.

As for the second sample, the student has failed to determine the volume of thrown water when the container is tilted. Rather than determining the volume of triangular prism, based on the explanation, the student assumed that the height for the tilted container is 15 cm. Therefore, the second step of the student was multiplying 8, 15, and 36.

Based on Figure 3, less than 10% of primary students were able to work on the problem correctly. We provide a sample of students' correct answer.

Sample of students' correct answer.

V container = 36cm×17cm×8cm=4896cm^3

V tilted = ½×8cm×8cm×15cm=480cm^3

Thus, the thrown water is=480cm^3 (B3)

The correct answer sample is illustrating that the student came up with the correct answer and able to solve the problem with correct strategies or procedures.

Junior High Level

From 30 items, there were 8, 14, and 8 questions for HOTS, MOTS, and LOTS, respectively. In other words, the number for MOTS items is the greatest of all. According to the analysis, it was found that the students' average score is very low. To add, the maximum score for junior

high level is lower than the maximum score on primary level. The statistical result for students in junior high school is displayed in Table 6.

Table 6: Statistical Data on Junior High Level

Statistical Data	Score (maximum of 100)
Maximum	73.33
Minimum	0
Average	30.85

Figure 5 presents a pie chart on the percentage of correct answer types. We found that most students who answered correctly on essay questions were able to explain in an appropriate strategies or procedures (B3). Based on the findings, the number of students who were able to determine the correct final answer but the strategies could not be identified as these were mysterious or undefined (B2), was at the very least. There is a slight difference among correct type B0, B1, and B2 compared to correct type B3.

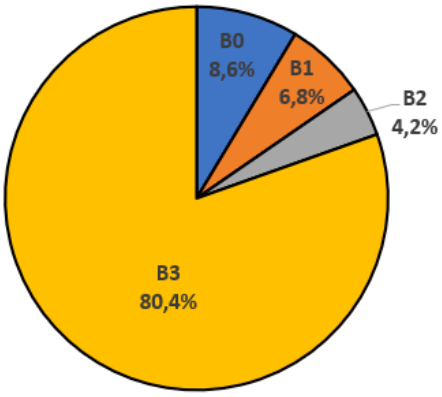

Figure 5: The percentage for each correct answer type on junior high level

As for the incorrect answers, we also determined its type of error illustrated by Figure 6. According to findings, most junior high students were unable to provide answer or they provided invalid answer (E6). Another error type that usually occurred is that students' strategy could not be identified due to limited explanation (E5). The very least type of error occurred was encoding (E4).

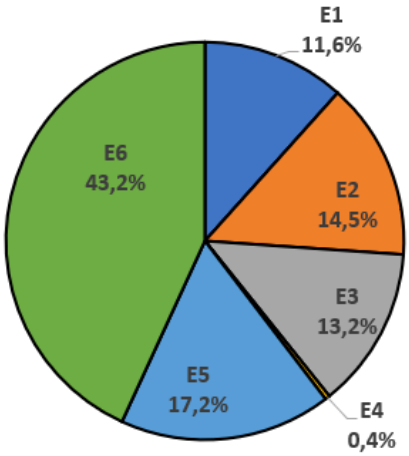

Figure 6: The percentage for each error type on junior high level

According to statistical analysis, the correct and error type for HOTS items is in line with findings the percentage for correct and error type on essay questions. For HOTS items on essay questions, most students were fall into error type E6 (42.8%) and correct type B3 (86.8%).

HOTS Items for Junior High Level

Figure 7 illustrates the percentage of students' answer for HOTS items. For question 23, none of junior high students were able to answer the problem. In addition, there are two items, following question 23, which is more than 90% of students failed to answer, those are question 22 and question 30. Compared to primary school HOTS essay questions, there is one HOTS item in junior high level which can be solved by more than a half of students grade 8 in MaRWA test, that is question 25.

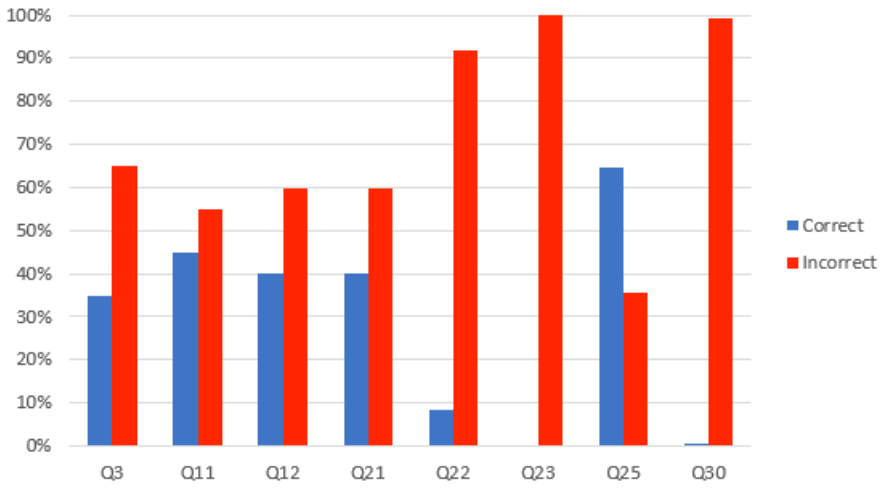

Figure 7: The percentage of junior high students' answer for HOTS items

To illustrate, we provide question 30 and samples of students' answer in Figure 8.

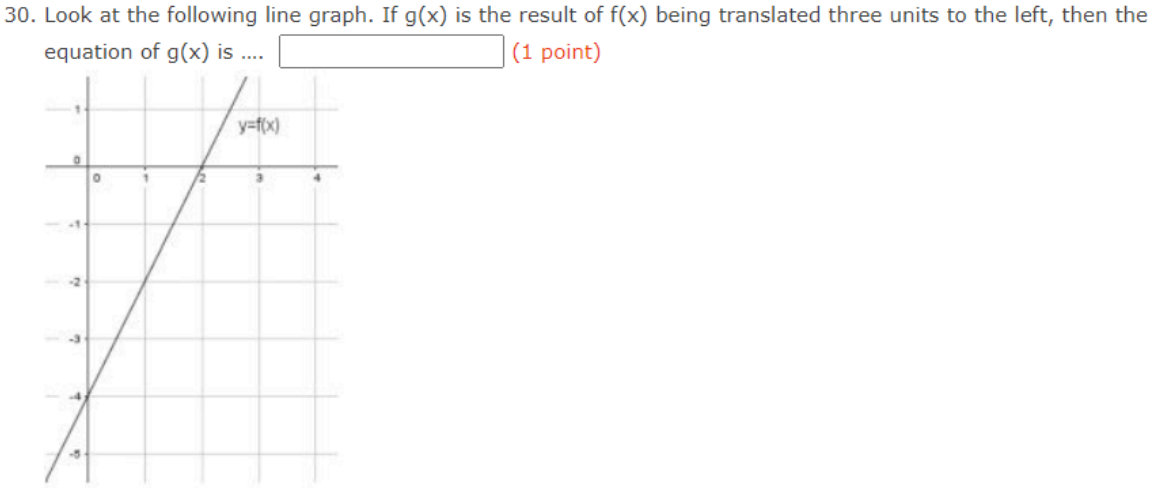

Figure 8: A sample of HOTS item for junior high level, question 30

Samples of students' error:

- Moving 3 units to the left means the x becomes -1 (E1)
- Because it intersects both (E5)

Above are two samples of error answers in different types. For the first sample, we generate that the student was failed to comprehend information provided in the problem (E1) as it simply determined the intersection between *g(x)* and *x* axis. This student was probably not aware that translating a line affects the line equation. For the second error sample, the student has failed to provide clearer information in solving the problem (E5). Indeed, the first line and the translated line intersect both *x* axis and *y* axis. However, this student wrote a limited explanation which could not be comprehend.

two samples are illustrating the error type of comprehending the problem (E1). The problem is asking for amount of water that is thrown out of the container when it is tilted. However, the first sample provides us an information that the student was unable to use information to solve the problem. Rather than determining the volume of a prism then subtracting it with the amount of thrown water, this student multiplied all numbers on the question. The student has failed to comprehend in determining the volume of a prism.

As for the second sample, the student has failed to determine the volume of thrown water when the container is tilted. Rather than determining the volume of triangular prism, based on the explanation, the student assumed that the height for the tilted container is 15 cm. Therefore, the second step of the student was multiplying 8, 15, and 36.

Based on Figure 7, less than 1% of junior high students were able to answer the problem correctly. In other words, there was only one correct answer among all. Here, we provide the only correct answer for question 30 MaRWA 2020.

Sample of students' correct answer.

> gy will pass through -1,0 and -3,-4 = a line equation which passes through these points are y=2x+2 (B3)

The correct answer sample is depicting the student's ability in comprehending information on question 30. The student translated the line and it knew that when the line moved three units to the left, it intersects point (-1,0) and (-3,-4). Even though, it did not provide the process on determining the line equation, the student knew that a line equation which passes through the two points is $y=2x+2$, which is the correct answer (B3).

Senior High Level

Similar to the other two levels, among 30 items on senior high level, there were 8, 14, and 8 questions for HOTS, MOTS, and LOTS, respectively. In other words, the number for MOTS items is the greatest of all. According to the analysis, it was found that the students' average score is very low. In addition, the average score in senior high level is the lowest among all. The statistical result for students in senior high school is displayed in Table 7.

Table 7: Statistical Data on Senior High Level

Statistical Data	Score (maximum of 100)
Maximum	66.67
Minimum	0
Average	24.52

Figure 9 displays a pie chart on the percentage of correct answer types. According to the analysis, most students who answered correctly on essay questions were able to explain in an appropriate strategies or procedures (B3). Based on the findings, the number of students who were able to determine the correct final answer but provided incorrect strategies (B1), was at the very least. This is similar to the occurrence in junior high level. Compared to other levels, the percentage of students in senior high, who only wrote correct answer without providing any detailed explanation (B0), is the highest.

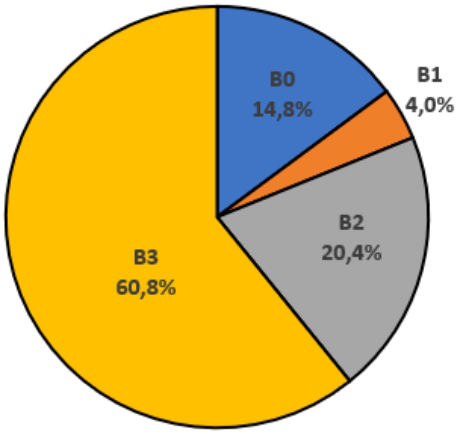

Figure 9: The percentage for each correct answer type on senior high level

According to findings on incorrect answers as displayed in Figure 10, most senior high students were unable to provide answer or they provided invalid answer (E6). It turns out that error type E6 is the most happening on three levels. Another error types that usually occurred are comprehension type (E1) and unknown type (E5). Encoding error type (E4) appears the least on all levels.

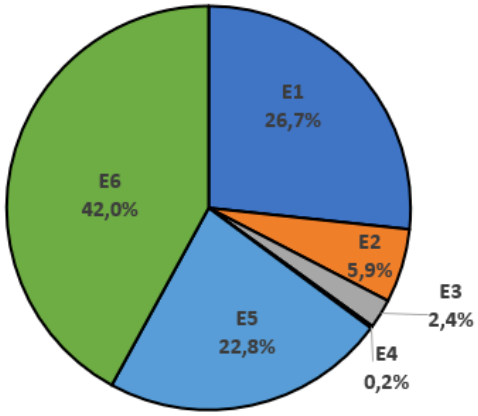

Figure 10: The percentage for each error type on senior high level

According to statistical analysis, the correct and error type for HOTS items is in line with findings the percentage for correct and error type on essay questions. For HOTS items on essay questions, most students were fall into error type E6 (38.9%) and correct type B3 (53.1 %).

HOTS Items Senior High Level

Figure 11 illustrates the percentage of students' answer for HOTS items. Answers for all HOTS items are mostly incorrect. It is similar to the finding on the primary level HOTS items. In specific, more than 70% of senior high students answered incorrectly for all HOTS items. Question 29 got the highest incorrect answer percentage among others, followed by question 27.

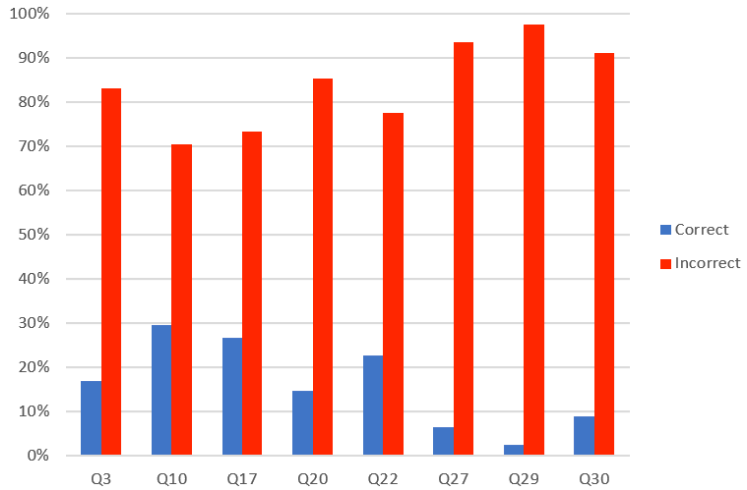

Figure 11: The percentage of senior high students' answer for HOTS items

To illustrate, we provide question 30 and samples of students' answer in Figure 12.

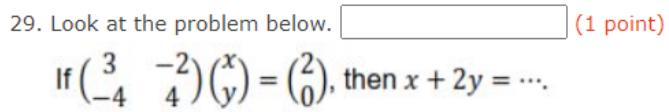

Figure 12: A sample of HOTS item for senior high level, question 29

Samples of students' error:

- (3 -2) (x) = (2)
  (-4 4) (y) = (0)

  3x – 2x = 2
  x = 2

  -4y + 4y = 0
  y = 0 (E2)
- If (3 -2) (x) then x+2y is 16 (E5)
  -4 4 (y)

Above are two samples of error answers in different types. For the first sample, the student was failed to use a proper procedure to solve the problem, which fell into error type transformation (E2). Rather than working on matrix linear equation system, $3x - 2y = 2$ and $-4x + 4y = 0$, the student directly subtracted the first row to determine the value of x and subtracted the second row to determine the value of y. In other words, the student was unable to utilised linear equation system on matrix and just proceeded with the calculation.

The second student noted down the given information, that is the left side of the linear equation system. However, the student it was obscure how it came up with 16. As there was no detailed information, this means that the student fell into type error unknown (E5), in which error type cannot be identified as the strategy used by the student provided limited information.

From the three levels, we found that the biggest portion and the most common type of students correct answer for essay questions is B3, while the least type of students correct answer differs, either B1 or B2. To add, the most occurred common error is E6, followed by E5 and E1. In contrast, it is rare for error type E4 to happen. Each level has eight HOTS items. In general, the percentage of incorrect answers for HOTS items is higher. Almost on all levels, none of correct answer on HOTS items percentage is higher than the incorrect answer, except for junior high level for question 25.

Conclusion

HOTS is one of parameters to measure students' ability in using critical thinking during problem solving and flexibilities to choose a strategy, which also becomes the reflection of mathematics learning. MaRWA, a diagnostic test assessing students' readiness in learning mathematics, found that the average score of students on all levels across Southeast Asia in 2020 was very low. Most students struggled to work on HOTS problems and tended to do error type E6, in which they left the answer blank or provided invalid final answer without providing any strategy or reasoning in detail. This can happen when students are not familiar in utilising their critical thinking skills. This indicates that the mathematics teaching and learning has not provide rooms for students to use their prior knowledge in solving problems. To respond to students' struggles, we encourage teachers to adjust teaching and learning into students-centred. It is also important to give students an opportunity to have discussion with their peers. When students having different strategies in working on the problem, the teacher can facilitate students by asking questions or to bridge one strategy to another to close the gap of conceptual construction. We also would recommend institution which caters teacher professional development to design a course related to students' HOTS and more pedagogy on allowing students to be flexible and open about their ideas.

References

Clements, M. A. (Ken). (1980). Analyzing Children's Errors on Written Mathematical Tasks. *Educational Studies in Mathematics*, *11*, 413–420.

Creswell, J. W. (2007). *Qualitative Inquiry and Research Design Choosing Among Five Approaches*. SAGE Publications Inc.

Eivers, E. (2010). PISA: Issues in implementation and interpretation. *Irish Journal of Education*, *38*, 94–118. https://www.jstor.org/stable/20789130

Forestier, K., & Adamson, B. (2017). A critique of PISA and what Jullien's plan might offer. *Compare: A Journal of Comparative and International Education*, *47*(3), 359–373. https://doi.org/10.1080/03057925.2016.1273095

Hopmann, S. T. (2008). No child, no school, no state left behind: Schooling in the age of accountability. *Journal of Curriculum Studies*, *40*(4), 417–456. https://doi.org/10.1080/00220270801989818

Lamichhane, B. R. (2018). Assessment Practices in Mathematics: Local to Global Contexts. *Saptagandaki Journal*, *9*(February), 1–16. https://doi.org/10.3126/sj.v9i0.20876

OECD. (2021). *Innovation in PISA: Frequently Asked Questions (FAQs) What are the PISA innovative domain assessments?* https://www.oecd.org/pisa/innovation/Frequently-Asked-Questions.pdf

Schleicher, A. (2019). PISA 2018: Insights and Interpretations. In *OECD*.

SEAQiM. (2012). *Working Paper 10: Concept Paper on Mathematics Regional Wide Assessment*.

Sjøberg, S. (2015). PISA and global educational governance - A critique of the project, its uses and implications. *Eurasia Journal of Mathematics, Science and Technology Education*, *11*(1), 111–127. https://doi.org/10.12973/eurasia.2015.1310a

White, A. L. (2017). Mathematics Literacy: An Australian Perspective and Experience. *Southeast Asian Mathematics Education Journal*, *7*(1), 31–42. https://doi.org/10.46517/seamej.v7i1.43

White, A. L. (2018). Australian Mathematics Literacy Perspectives: Assessment. *Southeast Asian Mathematics Education Journal*, *8*(1), 69–82. https://doi.org/10.46517/seamej.v8i1.65

Wijaya, A., van den Heuvel-Panhuizen, M., Doorman, M., & Robitzsch, A. (2014). Difficulties in solving context-based PISA mathematics tasks: An analysis of students' errors. *Mathematics Enthusiast*, *11*(3), 555–584. https://doi.org/10.54870/1551-3440.1317

Yin, R. K. (2003). *Case Study Research: Design and Methods*. SAGE Publications.

Zhao, Y. (2020). Two decades of havoc: A synthesis of criticism against PISA. *Journal of Educational Change*, *21*(2), 245–266. https://doi.org/10.1007/s10833-019-09367-x

Contact email: tyaswirani@staff.qitepinmath.org