2026, 7(1), 37–51

Journal of Didactic Mathematics
https://doi.org/10.34007/jdm.v7i1.3143

How computational thinking is represented in mathematics assessments: Insights from Indonesia, Malaysia, Singapore and their alignment with PISA

Ginda M.A. Siregar*, [a], Roni Priyanda[b]
[a,b]Universitas Samudra, Langsa, Aceh, Indonesia, 24416

Abstract.
Computational Thinking (CT) has emerged as an essential competency in twenty-first-century education and is increasingly integrated into mathematics curricula and assessment systems globally. Despite its growing significance, limited studies have examined how CT dimensions are represented in mathematics assessments across countries and how these dimensions align with the characteristics of the Programme for International Student Assessment (PISA). This study employed a qualitative approach using comparative content analysis. The data consisted of official junior secondary mathematics assessment documents administered in 2025 in Indonesia, Malaysia, and Singapore, which were selected through purposive sampling and analyzed using a coding framework encompassing decomposition, pattern recognition, abstraction, algorithmic thinking, and evaluation. The findings reveal substantial differences in the distribution of CT dimensions across countries. Indonesia predominantly emphasizes Decomposition (33.3%), reflecting a stronger focus on contextual understanding and information extraction. Malaysia places greater emphasis on Algorithmic Thinking (26.4%) and Abstraction (26.2%), indicating a focus on procedural fluency, mathematical modeling, and systematic problem-solving. Singapore demonstrates the highest proportion of Evaluation (15.4%) alongside strong representation of Abstraction (25.6%), highlighting formal reasoning, validation, and justification processes. Comparison with PISA tasks indicates that the assessments of the three countries share similarities in contextual problem-solving, data interpretation, and mathematical reasoning; however, they provide limited opportunities for complex system modeling, causal reasoning, simulation, and evidence-based argumentation.

Keywords:
Computational thinking; assessment analysis; examination questions; junior secondary school; comparative study

How to cite:
Siregar, G. M. A., & Priyanda, R. (2026). How computational thinking is represented in mathematics assessments: Insights from Indonesia, Malaysia, Singapore and their alignment with PISA. *Journal of Didactic Mathematics*, *7*(1), 37–51. https://doi.org/10.34007/jdm.v7i1.3143

INTRODUCTION

The digital transformation, characterized by advancements in artificial intelligence, cloud computing, big data, and automation, has redefined the competencies demanded of students in the 21st century. In this context, Computational Thinking (CT) is increasingly recognized as a fundamental competency that complements literacy, numeracy, critical thinking, and problem-solving. This is because it enables students to comprehend problems systematically through the processes of decomposition, pattern recognition, abstraction, and the design of solutions that can be effectively implemented (Falloon, 2024). Furthermore, CT not only supports logical and analytical thinking skills but also contributes to enhancing students' ability to solve intricate

* Corresponding author.
E-mail address: ginda.mas@unsam.ac.id

Received: 21 March 2026; Revised: 10 April 2026; Accepted: 28 April 2026
2721-5601/ © 2026 The Author(s). Published by Mahesa Research Center
This is an open access article under the CC BY license (https://creativecommons.org/licenses/by/4.0/)

problems, fosters creativity, and prepares them to confront the challenges of the digital society and the future workforce (Tikva & Tambouris, 2021). Numerous countries, international educational organizations, and PISA have identified CT as one of the key competencies in curriculum reform and educational assessment systems. Consequently, its integration into learning has become increasingly crucial for preparing an adaptable and innovative generation in the era of digital transformation (Bonilla-Castañeda et al., 2024).

The concept of Computer-Based Teaching (CT) was first widely introduced by Wing (2006) as the ability to formulate problems and devise solutions that can be effectively implemented by both humans and computers. In the context of education, CT generally comprises five main dimensions: decomposition, pattern recognition, abstraction, algorithmic thinking, and evaluation. Decomposition facilitates the breakdown of complex mathematical problems into simpler components, enabling easier analysis. Pattern recognition enables students to identify regularities, relationships, or structures that emerge in data and mathematical patterns. Abstraction aids students in focusing on pertinent information while disregarding extraneous details. Algorithmic thinking encourages students to formulate solution steps in a logical, systematic, and efficient manner. Evaluation plays a role in assessing the accuracy, effectiveness, and efficiency of the resulting solutions. The integration of these five dimensions into mathematics education has been demonstrated to support the development of mathematical reasoning, problem-solving, and higher-order thinking skills. Students are not only focused on the final result but also on the thought processes employed to arrive at a solution. This approach has also been applied to science and computer science education.

One method to assess the implementation of CT within the curriculum is through an assessment system capable of quantifying the extent to which these competencies have been developed in students. Assessments serve as a representation of the competencies anticipated by the curriculum and the education system (El-Hamamsy et al., 2024). An analysis of test items can provide insights into the extent to which the dimensions of CT—namely decomposition, pattern recognition, abstraction, algorithmic thinking, and evaluation—have been incorporated into educational assessment practices. As a comparative point of reference, cross-national comparative studies can be employed to identify variations in educational policy orientations and the characteristics of the assessments implemented. Singapore's mathematics education places a strong emphasis on abstraction, algorithms, and logical thinking as the foundation of CT (Ministry of Education Singapore, 2024). Malaysia, through the Standard Secondary School Curriculum (KSSM), integrates STEM education, digital literacy, and computational thinking to address the needs of the Fourth Industrial Revolution (Bahagian Pembangunan Kurikulum, 2018). Conversely, Indonesia, through the Merdeka Curriculum, emphasizes strengthening 21st-century competencies, developing the Pancasila Student Profile, digital literacy, creativity, collaboration, and the integration of CT to cultivate problem-solving skills pertinent to real-world contexts (Aditomo et al., 2024).

The disparities in curriculum orientation have the potential to influence the characteristics of assessments employed, which in turn may impact student performance in international studies such as the Programme for International Student Assessment (PISA). Singapore emerged as the global leader in PISA 2022, boasting a mathematics score of 575 points. Malaysia achieved a score of 409 points, while Indonesia secured a score of 366 points, all of which fell short of the OECD average (OECD, 2023). While PISA performance is influenced by various factors, including the quality of instruction, teacher preparedness, and socioeconomic conditions, differences in curriculum orientation towards the development of higher-order thinking skills and critical thinking (CT) are believed to contribute to the performance gap among students in these three countries (El-Hamamsy et al., 2024). Over the past decade, research on CT has experienced rapid advancements, yet most studies have continued to focus on the development of learning models, the assessment of student abilities, and the integration of CT into the curriculum (Falloon, 2024; Liu et al., 2024). This study endeavors to analyze and compare the distribution of CT dimensions encompassed within mathematics assessments and establish a correlation between these

dimensions and item characteristics as well as PISA scores. The findings of this study are anticipated to enrich the field of CT assessment while simultaneously serving as a foundation for the development of curriculum policies and the design of learning evaluations that effectively support the enhancement of students' CT competencies within the secondary education system.

METHOD

Research Approach, Methods, and Design

This study employs a qualitative approach utilizing comparative content analysis. It employs a cross-national comparative study design that focuses on comparing the distribution of cognitive test (CT) dimensions in final exam documents for junior high school or equivalent levels. Content analysis was selected because it enables researchers to systematically interpret assessment documents to reveal the characteristics and trends of the CT dimensions measured in each education system (Krippendorff, 2019).

The CT analysis framework was adapted from concepts proposed by (Wing, 2006) and (Brennan & Resnick, 2012). These concepts include decomposition, pattern recognition, abstraction, algorithmic thinking, and evaluation. This adaptation was made to identify and compare the CT dimensions contained in junior high school graduation exam questions in Indonesia or their equivalents in tests for 15-year-old students in Malaysia and Singapore, and to relate them to the Program for International Student Assessment (PISA) test.

Research Subject

The research subjects were selected using purposive sampling, specifically junior high school-level mathematics exam documents officially published by the educational authorities of each country in 2025. In Indonesia, mathematics assessment is conducted through an assessment organized by educational institutions, namely the Minimum Competency Assessment (AKM). In Malaysia, mathematics assessment at the end of Form 3 is conducted through the School-Based Assessment, which evaluates students' mastery on an ongoing basis through exams, assignments, and projects. In Singapore, mathematics assessment from Secondary 1 through Secondary 3 is conducted through school-based assessment, while at the end of Secondary 4, students take the Singapore-Cambridge GCE O-Level Mathematics examination, which measures competencies in the areas of Numbers and Algebra, and Geometry.

Instruments, Data Collection and Analysis Techniques

Data were collected through a documentary study of exam question papers obtained from official government sources or relevant educational institutions. The research subjects serving as data sources were test items classified by country, including the name of the test, the administration method (computer-based test (CBT) or paper-based test (PBT)), and the question types, which may include Multiple Choice (MC), Short Answer (SA), Multiple Response (MR), True/False (T/F Q), or Matching Questions (Match); the test-developing agency; and the test duration are summarized for each country in the following Table 1.

Table 1. Descriptive summary of the mathematics assessment

Country	Test Name	Type and Duration	Question Type and Number	Test Developer Indonesia
Indonesia	Junior High School Academic Ability Test (TKA) / Junior High School Numeracy Assessment	CBT in 105 min	36 questions: MC = 9, SA = 9, MR = 6, T/F = 9, matching = 3	Ministry of Education through the Center for Educational Assessment
Malaysia	Soalan tk 3 malaysia trengganu.pdf; the cover states SPM 2025 Learning Development Module, MPP	PBT in 1st paper=150 min, dan 2nd	57 questions: MC = 40, SA = 17	Ministry of Education of Malaysia, Terengganu State

Country	Test Name	Type and Duration	Question Type and Number	Test Developer Indonesia
	3, Mathematics Paper 1 and Paper 2	paper= 90 min		Education Department
Singapore	E-Math - Sec 3 - SA 2 - 2023 - Bukit Merah Sec.pdf; Mathematics Paper 1 & Paper 2, End-of-Year Examination 2023, Secondary 3 Express	PBT 1st paper=120 min, dan 2nd paper= 90 min	27 questions: Paper 1 = 21 questions, Paper 2 = 6 questions; predominantly written questions	Bukit Merah Secondary School

All test items were subsequently coded based on CT dimension indicators through an interpretation grounded in the CT dimensions. These dimensions encompass operational indicators for identifying the presence of decomposition, pattern recognition, abstraction, algorithmic thinking, and evaluation within each test item. The primary research instrument employed was the researcher's own interpretation (Creswell, 2014; Lincoln et al., 1985). The keywords utilized during the open coding process to identify the sub-dimensions and CT dimensions for each test item from each country are presented in Table 2.

Table 2. Coding framework for identifying computational thinking dimension

CT Dimension	Sub-dimension	Open Coding Keywords
Decomposition	Identifying information from stimuli	given information; table; graph; diagram; figure; image; data; statement; context; scenario
	Selecting relevant data	value; variable; measurement; quantity; distance; speed; volume; cost; temperature; probability; scale
	Breaking down problems into smaller parts	sub-question; part (a/b/c); component; step; parameter; formula; category; element
	Integrating multiple information sources	combine information; relate data; use information; compare representations; connect graph and table; interpret multiple sources
Pattern Recognition	Identifying numerical patterns	sequence; pattern; trend; relationship; frequency; difference; repetition; nth term
	Comparing quantities	greater than; less than; maximum; minimum; range; ratio; proportion; comparison
	Identifying statistical patterns	mean; median; mode; variance; distribution; histogram; scatter plot; frequency table
	Identifying spatial or geometric patterns	shape; transformation; symmetry; tessellation; coordinate; orientation; scale drawing; similar figures
Abstraction	Extracting essential information	relevant information; requirement; condition; constraint; key information; necessary data
	Transforming context into mathematical models	equation; expression; function; variable; inequality; matrix; formula; mathematical model
	Selecting representations or formulas	representation; graph; table; diagram; formula; notation; model; scale
	Simplifying contexts	simplify; express; factorise; convert; approximate; standardise; reduce complexity
Algorithmic Thinking	Determining solution steps	determine; calculate; find; solve; evaluate; complete; construct; write
	Applying procedures or formulas	apply formula; substitute; compute; perform operation; use method; execute procedure

CT Dimension	Sub-dimension	Open Coding Keywords
	Performing stepwise calculations	show working; step-by-step; sequential calculation; intermediate result; iterative process
	Developing systematic strategies	organise steps; choose strategy; construct graph; use previous answer; optimise solution
Evaluation	Verifying statements	verify; prove; validate; check; determine correctness; justify; confirm
	Assessing reasonableness of results	reasonable answer; estimate; significant figures; approximation; feasible result; accuracy
	Comparing alternatives	compare solutions; select best option; evaluate alternatives; determine efficiency; choose answer
	Validating processes and conclusions	explain reasoning; justify answer; draw conclusion; interpret result; provide evidence

Data analysis was conducted using descriptive and comparative content analysis. The frequency of occurrence for each cognitive test (CT) dimension was determined and converted into percentages to illustrate the distribution of CT dimensions across the three countries: Indonesia, Malaysia, and Singapore. Subsequently, a comparative analysis was performed to identify variations in the dominance of CT dimensions and trends in cognitive demand levels reflected in exam questions. The implications for each country's PISA performance were analyzed by examining several aspects of similarity to the sample questions provided in the "PISA 2025 Science Framework" document. The rationale for the inclusion of CT dimensions is presented in Table 3.

Table 3. Rasionalitas dimensi CT pada assasment PISA 2025

No	PISA 2025 Theme	Relevance to Mathematics Core Concepts
1	Greenhouse	Decomposition: Extracting information from a text and two graphs. Pattern Recognition: Identifying trends in CO_2 and temperature. Abstraction: Identifying the key variables in the greenhouse effect phenomenon. Evaluation: Assessing whether the datasupports the conclusion.
2	Smoking	Decomposition: This involves systematically analyzing research data, claims, and health information. Pattern Recognition: This entails identifying the correlation between cigarette consumption and disease risk. Abstraction: This process involves transforming health information into variable relationships. Evaluation: This step assesses the validity and scientific rigor of evidence and claims.
3	Zeer Pot	Decomposition: Identifying the components of the device and the variables that influence temperature. Abstraction: Modeling the relationships between sand depth, humidity, temperature, and food quantity. Algorithmic Thinking: Testing parameter combinations to achieve specific outcomes. Evaluation: Determining the most optimal conditions.
4	Running in Hot Weather	Decomposition: Analyzing the variables of temperature, humidity, hydration status, and physical risk. Abstraction: Understanding the thresholds for dehydration and heatstroke as a decision-making model. Algorithmic Thinking: Running simulations based on parameter changes. Evaluation: Assessing whether conditions are safe or risky based on data.
5	Data Dinosaurs	Decomposition: Analyzing dinosaur attributes and data presented in graphs. Pattern Recognition: Identifying data distribution patterns and relationships between variables. Abstraction: Simplifying data into a scatter plot representation. Evaluation: Assessing the validity of conclusions based on the data.
6	Top Predators	Decomposition: Analyzing diagrams that depict the relationships between organisms. Pattern Recognition: Identifying patterns of direct and indirect relationships. Abstraction: Comprehending ecosystems as cause-and-effect network models. Evaluation: Assessing the impact of modifications to a single component on the system.

RESULTS AND DISCUSSION

Distribution of Computational Thinking Dimensions

The distribution of cognitive test dimensions (CTD) was analyzed through a coding process applied to all junior high school-level mathematics test items from Indonesia, Malaysia, and Singapore. The identification process involved examining the cognitive demands and solution characteristics required for each item, followed by calculating the frequency of occurrence for each CTD. Subsequently, these frequencies were converted into percentages to illustrate the proportion of each CTD in the total number of items analyzed. The frequencies of CTD for each country are presented in Table 4.

Table 4. Frequency of occurrence of CT dimensions by country

Sub-dimension	Count (open coding)			Description
	Id	My	Sg	
Decomposition				
Identifying information from stimuli/diagrams/tables	36	57	27	All questions require the identification of initial information from text, tables, graphs, diagrams, or question instructions.
Sorting relevant data	27	53	27	Very dominant in Malaysia and Singapore because almost all questions require sorting numerical/symbolic data.
Breaking down problems into smaller parts	16	39	19	Higher in Malaysia because many questions on Paper 1 and Paper 2 contain multiple computational or procedural components.
Connecting multiple pieces of information	9	35	15	Appears when students must combine two or more pieces of information, such as a table and conditions, a graph and a formula, or context and calculations.
Pattern Recognition				
Recognizing number/data patterns	11	17	9	Appears in number patterns, graphs, tables, sequences, functions, and data changes.
Comparing quantities	12	27	10	Appears frequently in Malaysia because questions involve comparing values, graphs, finance, statistics, and geometry.
Recognizing statistical/data patterns	6	7	2	Indonesia and Malaysia include more statistical data, probability, ogives, histograms, min, mode, median, and variance.
Recognizing spatial/geometric patterns	14	27	13	Dominant in Malaysia due to the prevalence of geometric figures, transformations, tessellations, graphs, and coordinates.
Abstraction				
Extracting Key Information	10	57	27	In Malaysia and Singapore, nearly all questions require students to extract key information from formal mathematical expressions.
Transforming Context into Mathematical Models	18	48	25	This occurs when the context of a story, table, graph, or diagram is transformed into equations, probabilities, areas, volumes, matrices, functions, or computational models.
Selecting Representations/Formulas	15	50	25	High in Malaysia and Singapore because many problems use formulas, graphs,

Sub-dimension	Count (open coding)			Description
	Id	My	Sg	
Decomposition				
				matrices, functions, transformations, geometry, and statistics.
Simplifying the Context	8	31	13	Occurs when students must identify the core of the problem's conditions and discard unnecessary information.
Algorithmic Thinking				
Determining the sequence of solution steps	16	57	27	Very dominant across all test papers, particularly in Malaysia and Singapore, which require systematic mathematical procedures.
Applying procedures/formulas	18	53	25	Appears in problems involving calculations, graphs, algebra, probability, matrices, geometry, statistics, and functions.
Performing step-by-step calculations	11	42	18	Malaysia scored the highest because many questions on Papers 1 and 2 required step-by-step operations.
Developing systematic strategies	9	36	15	Appears in questions requiring a sequence of strategies, not just single operations.
Evaluation				
Checking the truth of statements/representations	9	14	8	Indonesia is strong in true-false questions; Malaysia and Singapore excel in questions involving the validation of statements, graphical representations, and proofs.
Assessing the plausibility of results	5	27	10	This skill appears frequently in Malaysia because some questions require selecting the most plausible answer based on calculations or context.
Comparing alternative solutions	9	17	9	Appears in questions involving alternative choices, combinations, optimal values, or the most appropriate answer option.
Validating processes and conclusions	5	17	27	Singapore scored the highest because the exam instructions emphasized working out the problem, the accuracy of steps, and justification of answers.

Based on the tabulation results, Algorithmic Thinking emerges as the most prominent CT dimension across all three test versions. This dimension achieved the highest score due to the majority of math questions necessitating students to adhere to systematic problem-solving procedures, apply formulas, execute step-by-step calculations, and ascertain problem-solving strategies. In the Indonesian test, Algorithmic Thinking is evident in questions that require students to calculate total calories, fat content, probability, capacity, area, volume, cost, and price. In the Malaysian test, this dimension exhibits even greater dominance, as the questions on Papers 1 and 2 predominantly involve formal procedures such as algebra, graphs, geometry, matrices, functions, probability, statistics, and financial calculations. In the Singapore test, Algorithmic Thinking is also highly prominent, as the question structure necessitates students to demonstrate their work, solve equations, create graphs, calculate length, area, and volume, and elucidate the solution process.

Conversely, Pattern Recognition, particularly its sub-dimension of statistical/data pattern recognition, appears the least overall. This is attributed to the fact that not all questions explicitly require students to identify patterns, trends, or regularities in the data. In the Indonesian test,

Pattern Recognition is evident in questions pertaining to food reviews, game odds, light bulb quality control, house numbering patterns, and several data-related inquiries. In the Malaysian test, this dimension is discernible in number patterns, graphs, transformations, ogives, histograms, statistics, and geometry. In the Singapore test, Pattern Recognition is present in sequences, functions, graphs, scales, similarity, and relationships between variables. However, in comparison to Algorithmic Thinking and Abstraction, its frequency is lower, as many questions emphasize the application of procedures rather than pattern recognition.

There are numerous questions that incorporate multiple cognitive dimensions. This is anticipated because mathematical questions, particularly those based on stimuli, generally assess more than one cognitive process. For instance, a single question derived from a nutrition table in the Indonesian test paper may involve Decomposition when students sort information from the table, Abstraction when they convert the information into a mathematical model, Algorithmic Thinking when they calculate the total content, and Evaluation when they determine whether a statement is true or false. In the Malaysian test, questions involving graphs, matrices, statistics, and finance also frequently incorporate multiple dimensions simultaneously, as students must interpret data, select formulas, perform calculations, and then evaluate the correct answer. In the Singaporean test, structured questions such as those in geometry, functions, graphs, and trigonometry also demonstrate a combination of cognitive dimensions, as students must comprehend information, construct mathematical models, execute procedures, and validate results through step-by-step work.

In summary, the distribution of cognitive dimensions can be categorized into basic cognitive dimensions and advanced cognitive dimensions. Basic cognitive dimensions are evident in the Decomposition dimension and parts of Pattern Recognition—specifically when students are asked to identify information, sort data, read tables or graphs, recognize simple patterns, compare quantities, and understand basic relationships between variables. Advanced cognitive dimensions are evident in the Abstraction, Algorithmic Thinking, and Evaluation dimensions, specifically when students must construct mathematical models from a context, select the appropriate formula or representation, outline solution steps, perform step-by-step calculations, compare alternative solutions, and validate results. The table illustrates that all three test booklets not only include basic cognitive dimensions but also strongly emphasize advanced cognitive dimensions, particularly in the Malaysian and Singaporean booklets, which rely more heavily on formal, procedural, and structured problems. The Indonesian booklet also includes advanced cognitive dimensions, but it appears in the form of contextual numeracy and stimulus-based statement evaluation.

Computational Thinking Dimensions Across Indonesia, Malaysia, and Singapore

Upon comparing the proportions of each CT dimension across countries, it becomes evident that the three test items exhibit distinct characteristics and emphases. In the Indonesian test item, Decomposition emerges as the most prominent dimension, accounting for 33.3% of the total. This is followed by Algorithmic Thinking at 20.5%, Abstraction at 19.3%, Pattern Recognition at 16.3%, and Evaluation at 10.6%. This suggests that the Indonesian test places a greater emphasis on students' ability to interpret the stimulus, discern key information, extract data from tables, graphs, and images, and comprehend the problem's context prior to executing calculations. Conversely, in the Malaysian test, Algorithmic Thinking holds the highest proportion at 26.4%, followed by Abstraction at 26.2%, Decomposition at 25.9%, Pattern Recognition at 11.0%, and Evaluation at 10.5%. This distribution indicates that the Malaysian test prioritizes problem-solving procedures, formula selection, mathematical modeling, and systematic calculations. Notably, in the Singapore test paper, Abstraction stands out as the most significant dimension, comprising 25.6% of the total. This is followed by Decomposition at 25.1%, Algorithmic Thinking at 24.2%, Evaluation at 15.4%, and Pattern Recognition at 9.7%. Consequently, the Singapore test demonstrates a stronger capacity to transform problems into formal mathematical representations, utilize symbolic representations, and validate the process and

outcomes of problem-solving. A comparative analysis of the proportions of each dimension across countries is presented in Figure 1.

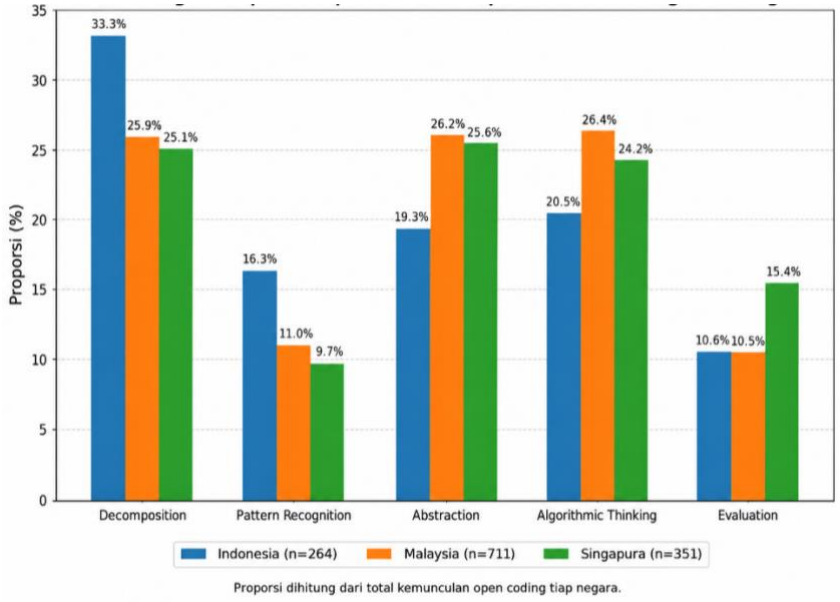


Figure 1. Distribution of computational thinking dimensions by country

When considering the dominance of each dimension, Decomposition exhibits the highest level in Indonesia, signifying the strong inclination of Indonesian test items toward decomposing information derived from contextual numeracy stimuli. Indonesia also achieved the highest score in Pattern Recognition, as numerous questions necessitated students to discern data patterns, compare quantities, identify trends, and establish connections between information presented in tables or graphs. Malaysia, on the other hand, attained the highest score in Abstraction, albeit with a narrow margin compared to Singapore, indicating a robust emphasis on transforming problem information into models, equations, formulas, or mathematical representations. Malaysia also demonstrated the highest score in Algorithmic Thinking, suggesting that the Malaysian test stands out most in terms of procedures, solution steps, and the application of mathematical algorithms. Conversely, Singapore secured the highest score in Evaluation, implying that the Singaporean test places a greater emphasis on requiring students to validate the accuracy of their processes, demonstrate their work, elucidate their answers, and ensure the precision of graphical, geometric, and algebraic representations. Consequently, Indonesia exhibits a stronger orientation towards comprehending stimuli and contextual problem-solving; Malaysia prioritizes procedures and formal mathematical applications; while Singapore emphasizes modeling, formal reasoning, and answer validation.

Each country's test papers demonstrate that a single question can encompass multiple dimensions of critical thinking (CT). Furthermore, the distribution of CT also reveals distinctions between basic CT and advanced CT. In the Indonesian test, basic CT is more prevalent, particularly Decomposition and Pattern Recognition. This is evident because many questions are stimulus-based and involve information sorting and pattern recognition in everyday contexts. In contrast, the Malaysian and Singaporean tests emphasize advanced CT, particularly Abstraction, Algorithmic Thinking, and Evaluation. Malaysia demonstrates advanced CT through a variety of procedures, formulas, and formal mathematical problem-solving. Singapore, on the other hand, demonstrates advanced CT through symbolic representation, proofs, working out solutions, and process validation. The gaps in CT distribution are most noticeable in Decomposition, which is highly prevalent in Indonesia; Algorithmic Thinking and Abstraction, which are strong in Malaysia; and Evaluation, which is most prominent in Singapore.

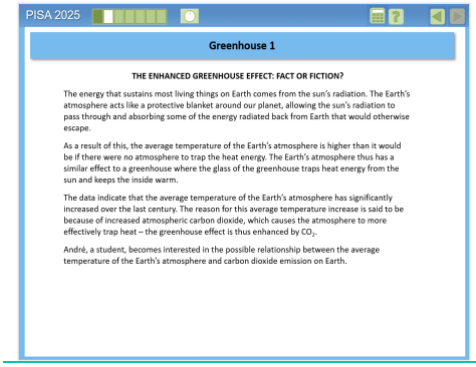
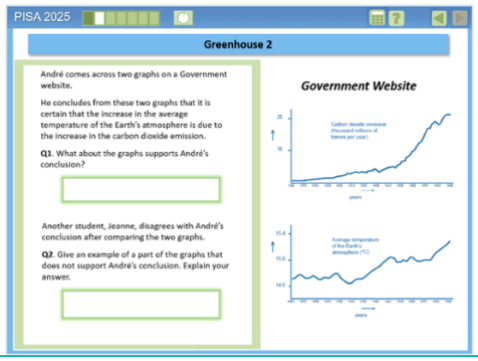
Implication to PISA

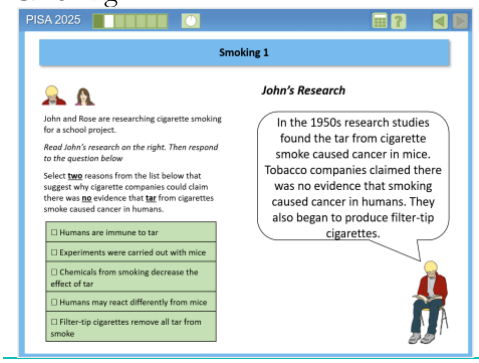
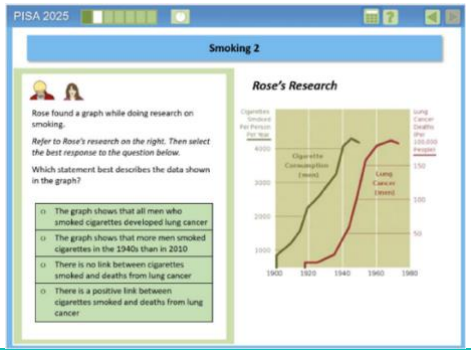
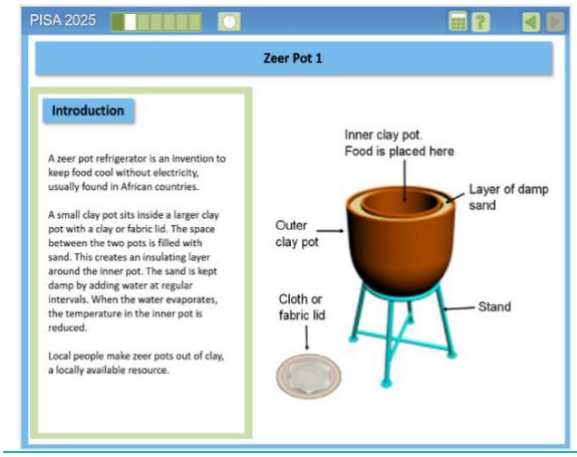
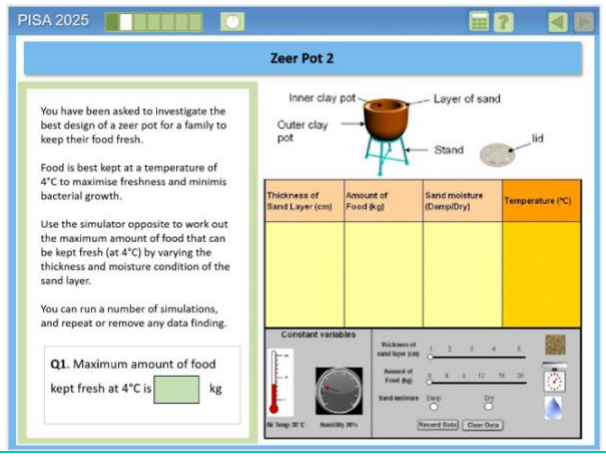
It is not accurate to generalize that the Grade 3/junior high school final exam from any one country is the most effective. Each test is tailored to the specific curricula, assessment objectives, grade levels, question formats, and evaluation traditions of its respective country. For instance, the Indonesian test is more aligned with stimulus-based numeracy assessment, while the Malaysian test is a formal mathematics exam comprising Paper 1 and Paper 2. The Singaporean test, on the other hand, reflects a school mathematics exam format that necessitates written solutions and process validation. Therefore, a comparison of the three tests should not be intended to determine which country is superior, but rather to identify the trends in the orientation of CT dimensions emphasized in each instrument.

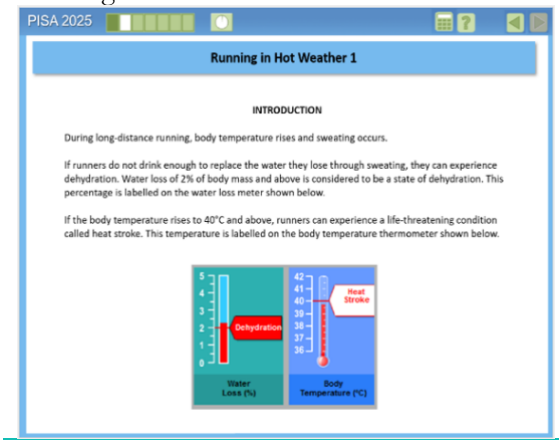
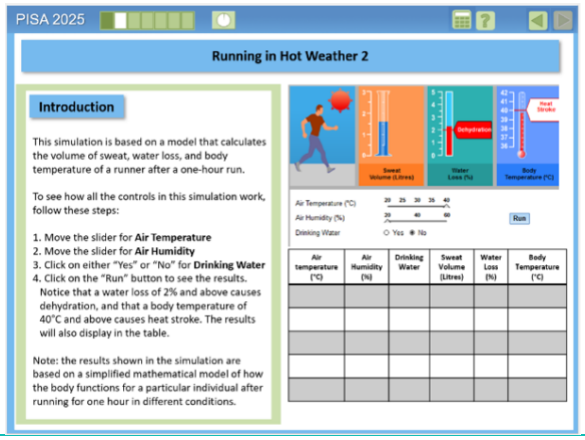
However, based on the PISA 2022 results, there are indications that assessment systems that place a stronger emphasis on abstraction, algorithmic thinking, evaluation, and problem-solving are associated with higher mathematics literacy performance. In PISA 2022, Singapore's math score reached 575, significantly surpassing the OECD average of 472; Malaysia scored 409, while Indonesia scored 366. Furthermore, the OECD reported that 41% of Singaporean students achieved Level 5 or 6 in mathematics—levels that require the ability to model complex situations and to select, compare, and evaluate problem-solving strategies. In contrast, in Indonesia, only 18% of students attained at least Level 2 in mathematics, while in Malaysia, 41% of students achieved at least Level 2; both figures remain below the OECD average of 69%.

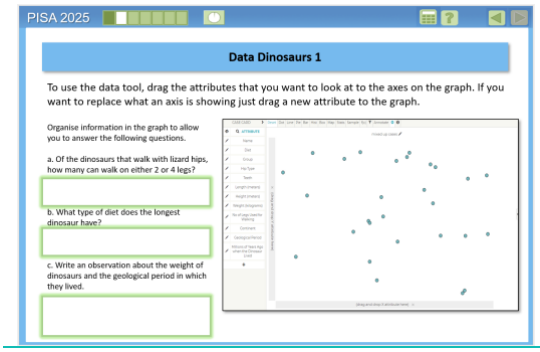
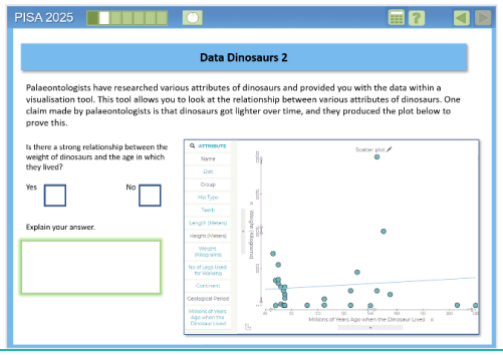
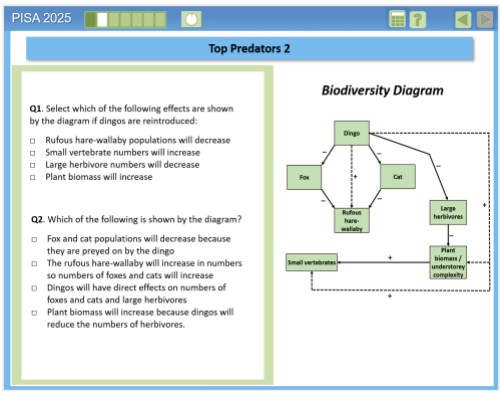
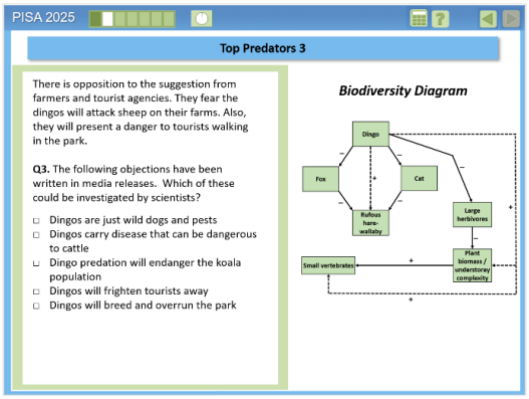
Several comparisons examine the alignment of PISA sample questions—as presented in the document "PISA 2025 Assessment and Analytical Framework"—with each country's assessments, as outlined in Table 5.

Table 5. Mapping of computational thinking dimensions in PISA mathematics problems

Dimensi CT	Indonesia	Malaysia	Singapore
Greenhouse				
Decomposition	Moderate – students read stimuli, data, graphs, and temperature/health contexts.	Moderate – students read diagrams, graphs, ogives, and data on changes.	Moderate–strong – students read graphs and structured information.
Pattern Recognition	Moderate – recognizing trends in graphs/data, particularly in quality control and food reviews.	Moderate – recognizing patterns in graphs, ogives, histograms, and changes in values.	Moderate – recognizing patterns of change in function/motion graphs.
Abstraction	Weak–moderate – not yet strong at transforming phenomena into cause-and-effect models.	Moderate – strong in mathematical representations, but not yet in modeling scientific phenomena.	Moderate–strong – closer to graphing and interpreting changes.
Algorithmic Thinking	Weak–moderate – primarily reading data, with few investigative procedures.	Moderate – includes graphical and computational procedures, but not	Moderate – includes procedures for reading and connecting graphs.

Dimensi CT	Indonesia	Malaysia	Singapore
		cause-and-effect simulations.	
Evaluation	Moderate – evaluating statements as consistent or inconsistent with the data.	Moderate – evaluating results from graphs/statistics, but lacking in scientific claims.	Moderate–strong – closer to justifying and validating answers.
Smoking			
Decomposition	Moderate–Strong – close to health stimuli, GGL, and disease risks.	Moderate – close to risk data, insurance, statistics, and graphs.	Moderate – reading information, graphs, and the context of the question.
Pattern Recognition	Moderate – recognizing patterns in health data and quality control.	Moderate–strong – strong in statistical patterns, graphs, ogifs, and plots.	Moderate – recognizing relationships between variables in graphs.
Abstraction	Moderate – converting consumption/risk data into simple decisions.	Moderate – converting data into calculation/statistical procedures.	Moderate–strong – converting context into mathematical models and reasoning.
Algorithmic Thinking	Weak–moderate – not dominant because there is more evaluation of statements.	Moderate – using graphical/statistical procedures.	Moderate – using step-by-step procedures and supporting calculations.
Evaluation	Moderate – assessing whether statements are true or false based on data.	Moderate – evaluating graphical/statistical information, but less so regarding scientific claims.	Moderate–Strong – stronger in explanations, reasoning, and validation.
Zeer Pot			

Dimensi CT	Indonesia	Malaysia	Singapore
Decomposition	Moderate–Strong – strong on problems involving tools, temperature, capacity, and water purifiers.	Moderate – interprets diagrams, tables, functions, and applied science contexts.	Moderate – interprets physical situations, graphs, and motion/fluid contexts.
Pattern Recognition	Moderate – recognizing patterns in temperature, size, capacity, and conditions.	Moderate – recognizing patterns in graphs, functions, and variable relationships.	Moderate – recognizing patterns of change in graphs.
Abstraction	Moderate–strong – transforming contexts involving tools and temperature into mathematical operations.	Moderate–strong – strong in modeling functions, graphs, variations, and formulas.	Moderate–Strong – Strong in modeling situations as graphs or representations.
Algorithmic Thinking	Moderate – Calculating capacity, size, temperature, and conditions.	Moderate–Strong – Strong in function procedures, variations, and formal calculations.	Moderate – Using graphical procedures and calculations.
Evaluation	Moderate – Assessing the adequacy of size or capacity based on conditions.	Moderate – evaluating procedural outcomes, but less so in investigative simulations.	Moderate–Strong – stronger in validating processes and outcomes.
Running in Hot Weather			
Decomposition	Moderate–Strong – close to issues of health, temperature, risk, and condition data.	Moderate – reading graphs, tables, statistics, and numerical conditions.	Moderate – reading graphs, conditions, and representations of motion.
Pattern Recognition	Moderate – recognizing patterns in temperature, risk, and data comparisons.	Moderate – recognizing patterns in graphs, statistics, and changes.	Moderate – recognizing patterns of change from graphs.
Abstraction	Moderate – converting temperature/risk information into numerical decisions.	Moderate–strong – converting conditions into functions, graphs, and computational models.	Moderate–strong – strong in modeling graphs and conditions.
Algorithmic Thinking	Moderate – using simple formulas/rules to calculate and select data.	Moderate–Strong – Strong in graphing procedures, statistics, and functions.	Moderate – Using step-by-step procedures to connect graphs and results.

Dimensi CT	Indonesia	Malaysia	Singapore
Evaluation	Moderate – Assessing safety/risk conditions based on data.	Moderate – Evaluating results from data/graphs, but weaker in risk simulation.	Moderate–Strong – Strong in process validation and justifying answers.
Data Dinosaurs			
Pattern Recognition	Moderate–Strong – recognizes data patterns, comparisons, and trends.	Moderate–Strong – closest due to proficiency with scatter plots, ogives, histograms, and statistical patterns.	Moderate – recognizes graph/function patterns, but not scatter plots.
Abstraction	Moderate – filters information from data and transforms it into conclusions.	Moderate–Strong – selecting statistical representations and computational models.	Moderate–Strong – modeling graphical relationships and symbolic representations.
Algorithmic Thinking	Moderate – calculating, comparing, and determining statements based on data.	Moderate–Strong – performing statistical procedures, graphing, and calculations.	In progress – running graph and function procedures.
Evaluation	Moderate – evaluates data- or graph-based statements.	Moderate – evaluates statistical results, but is not yet strong in interpreting scatter plots.	Moderate–strong – strong in process validation, though not a data tool.
Top Predators			
Decomposition	Moderate – reads stimuli and relationships between elements, but not in the form of ecosystem diagrams.	Moderate–Strong – closest to directed graphs, network graphs, and relationships.	Moderate – interprets diagrams/graphs and structured information.
Pattern Recognition	Moderate – recognizes patterns of relationships, probabilities, rules, and true/false statements.	Moderate–Strong – recognizes relationships between objects through graphs, sets, and diagrams.	Moderate – recognizes graphical patterns and representational relationships.

Dimensi CT	Indonesia	Malaysia	Singapore
Abstraction	Moderate – simplifying relationships/patterns into decisions.	Moderate–strong – understanding graphs as models of mathematical relationships.	Moderate–strong – transforming information into graphical models/representations.
Algorithmic Thinking	Weak–moderate – limited use of chained relational procedures.	Moderate – using graph, set, and relational procedures.	Moderate – using graph-based procedures and modeling.
Evaluation	Moderate – evaluating whether statements align with or deviate from the stimulus.	Moderate–strong – evaluating the truth of statements based on graphs/relationships.	Moderate–strong – strong in validation, working through problems, and explanation.

Upon analyzing the test items of each country and their resemblance to PISA items, it is evident that no single country's test items are unequivocally superior. Instead, a comprehensive approach to the dimensions of computational thinking is crucial for enhancing the quality of mathematics assessment. Instruments that exclusively focus on information sorting or computational procedures are inadequate in fostering high achievement as measured by PISA. Assessments must strengthen questions that necessitate students to construct mathematical models from context, select solution strategies, compare alternatives, verify the plausibility of results, and provide justifications. This approach appears to be more closely aligned with the characteristics of questions that demand advanced computational thinking, particularly abstraction and evaluation, which, according to the PISA 2025 results, are key attributes of high-performing students.

CONCLUSIONS

The concept of Computational Thinking (CT) has been incorporated into the mathematics assessment instruments of Indonesia, Malaysia, and Singapore, albeit with varying emphases contingent upon the distinct characteristics of each country's curriculum and assessment methodologies. Indonesia places a greater emphasis on Decomposition and Pattern Recognition through stimulus-based numeracy questions. Conversely, Malaysia prioritizes Abstraction and Algorithmic Thinking through formal and procedural mathematical questions. Singapore, on the other hand, emphasizes Abstraction, Algorithmic Thinking, and Evaluation through structured questions that necessitate modeling, problem-solving, and the validation of answers. A comparative analysis of PISA sample questions reveals that the questions from the aforementioned countries exhibit similarities in terms of reading data, graphs, patterns, and problem-solving. However, they do not yet fully align with the PISA framework, which places a greater emphasis on simulation, the evaluation of claims, cause-and-effect relationships, system modeling, and the utilization of data to support arguments. The findings of this study suggest the necessity of developing mathematics assessment designs that not only assess procedural skills but also enhance advanced CT, particularly abstraction, evaluation, modeling, solution validation, and data-driven decision-making. Test developers at the junior high school level should consider increasing the number of test items based on complex contexts, multiple representations, simulations, argumentation, and claim evaluation. This approach would better prepare students to meet the demands of international assessments such as PISA.

REFERENCES

Aditomo, A., Syahril, I., Yuliati, K., Suryani, N., Muhammad, H., & Handayani, F. (2024). *Kajian akademik Kurikulum Merdeka*. Pusat Kurikulum dan Pembelajaran, Badan Standar, Kurikulum, dan Asesmen Pendidikan, Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi.

Bahagian Pembangunan Kurikulum. (2018). *Dokumen standard kurikulum dan pentaksiran tingkatan 4 dan 5 matematik tambahan*. Kementerian Pendidikan Malaysia.

Bonilla-Castañeda, M. N., Villalba-Condori, K. O., Cardona-Reyes, H., Acra-Despradel, C., & Turner-Lam, K.-F. (2024). Global trends in computational thinking in curricula: A comparative review. *International Conference on Computers in Education*. https://doi.org/10.58459/icce.2024.5049

Brennan, K., & Resnick, M. (2012). New frameworks for studying and assessing the development of computational thinking. In *Proceedings of the 2012 Annual Meeting of the American Educational Research Association*. https://scratched.gse.harvard.edu/ct/files/AERA2012.pdf

Büscher, C. (2025). Differences in students' computational thinking activities when designing an algorithm for drawing plane figures. *International Journal of Science and Mathematics Education, 23*(2), 365–386. https://doi.org/10.1007/s10763-024-10465-3

Chytas, C., Van Borkulo, S. P., Drijvers, P., Barendsen, E., & Tolboom, J. L. J. (2024). Computational thinking in secondary mathematics education with GeoGebra: Insights from an intervention in calculus lessons. *Digital Experiences in Mathematics Education, 10*(2), 228–259. https://doi.org/10.1007/s40751-024-00141-0

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). SAGE Publications.

El-Hamamsy, L., Zapata-Cáceres, M., Martín-Barroso, E., Mondada, F., Zufferey, J. D., Bruno, B., & Román-González, M. (2024). The competent computational thinking test (cCTt): A valid, reliable and gender-fair test for longitudinal CT studies in grades 3–6. *arXiv*. https://doi.org/10.48550/arXiv.2305.19526

Falloon, G. (2024). Advancing young students' computational thinking: An investigation of structured curriculum in early years primary schooling. *Computers & Education, 216*, Article 105045. https://doi.org/10.1016/j.compedu.2024.105045

Krippendorff, K. (2019). *Content analysis: An introduction to its methodology* (4th ed.). SAGE Publications.

Lehmann, T. H. (2024). How current perspectives on algorithmic thinking can be applied to students' engagement in algorithmatizing tasks. *Mathematics Education Research Journal, 36*(3), 609–643. https://doi.org/10.1007/s13394-023-00462-0

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. SAGE Publications.

Liu, Z., Gearty, Z., Richard, E., Orrill, C. H., Kayumova, S., & Balasubramanian, R. (2024). Bringing computational thinking into classrooms: A systematic review on supporting teachers in integrating computational thinking into K–12 classrooms. *International Journal of STEM Education, 11*(1), 51. https://doi.org/10.1186/s40594-024-00510-6

Ministry of Education Singapore. (2024). *MOE Committee of Supply 2024*. https://www.moe.gov.sg/microsites/cos2024/

OECD. (2023). *PISA 2022 results (Volume I): The state of learning and equity in education*. OECD Publishing. https://doi.org/10.1787/53f23881-en

Tikva, C., & Tambouris, E. (2021). Mapping computational thinking through programming in K–12 education: A conceptual model based on a systematic literature review. *Computers & Education, 162*, Article 104083. https://doi.org/10.1016/j.compedu.2020.104083

Wing, J. M. (2006). Computational thinking. *Communications of the ACM, 49*(3), 33–35. https://doi.org/10.1145/1118178.1118215