ESTEEM Academic Journal
Vol. 17, August 2021, 11-22

Development of a Rubric to Assess Computational Thinking Skills among Primary School Students in Malaysia

Ling-Ling, Ung[1], Jane Labadin[2] and Saiful Nizam[3]
[1, 3]Faculty Computer & Mathematical Sciences, Universiti Teknologi MARA (UiTM), Locked Bag 71, 88997 Kota Kinabalu, Sabah, MALAYSIA
[2]Institute of Social Informatics and Technological Innovations, Jalan Datuk Mohammad Musa, 94300 Kota Samarahan, Sarawak, MALAYSIA
**corresponding author: [1]ungli720@uitm.edu.my*

ARTICLE HISTORY

Received
3 May 2021

Accepted
29 June 2021

Available online
4 August 2021

ABSTRACT

In the revised curriculum 2017, computational thinking skills have been integrated into the curriculum contents of all existing subjects in primary and secondary schools in Malaysia. The newly revised curriculum calls for an urgent need to prepare teachers to deliver computational thinking skills in the classroom and assess the teaching and learning outcomes of computational thinking skills. This paper reports the development of an assessment rubric, myCTRubric, designed to evaluate the teaching and learning outcomes of the newly integrated computational thinking skills among primary school students in Malaysia. An action research approach guided the design and development of myCTRubric. myCTRubric was reviewed and validated by a panel of computational thinking subject matter experts. Analyses of results show that myCTRubric is reliable and has strong content validity. This pioneering work is expected to lend insights into the teaching and learning practices of computational thinking in Malaysian classrooms and served as a practical guideline for future research.

Keywords: *Computational thinking skills, assessment, rubric, teaching and learning, primary school.*

1. INTRODUCTION

For many years, children's learning has focused on literacy and mathematics. Attention was then shifted to science and technology, integrating technologies into classrooms [1, 2]. As technologies evolve, young children's technology literacy has varied, leading researchers to question the relevance of the existing science and technology curricula to their knowledge development [3-5]. Researchers suggest that the current curriculum is moulding children to be end-users of technology by knowing how to use the technologies; however, they do not understand the development that impede their capabilities in thinking, learning, and creating [6, 7]. Children should design and develop technologies [8, 9], solve problems, and acquire interdisciplinary skills and knowledge [10, 11].

Malaysian children have been formally introduced to computer lessons in Primary 1 since 1970 [12]. The main objectives of the lessons are to expose learners to the functions available in computer applications. They were taught basic computer knowledge and applications. This can be observed from the Information, Communication, and Technology (ICT) syllabus and teachers' teaching modules from Kementerian Pendidikan Malaysia [13]. The situation changed

p-ISSN 1675-7939; e-ISSN 2289-4934

when the then Prime Minister of Malaysia announced integrating computational thinking (CT) skills into all subjects starting from2017 with Primary 1 and Form 1 students [14].

With the newly revised curriculum, there is an urgent need to equip schools, especially teachers, with CT skills in their daily classroom practices. The first step of integrating CT in the curriculum is by preparing the teachers to deliver the CT concepts in their daily teaching and learning (TL) practices. Professional developments courses are organised to raise teachers' understanding of the improved curriculum in preparing them with TL approaches that can be practised [14] in the classroom. However, the course content has not indicated any tools or techniques to measure CT TL outcomes.

2. LITERATURE REVIEW

2.1 Computational Thinking in Malaysia

Numerous researchers have agreed on the importance of mastering CT skills that CT is a must-have skill to survive in the 21st century, especially in today's workforce environment [15-17]. The revised curriculum incorporates CT skills in daily TL through brain-based teaching structures, inquiry-based approaches, and collaboration techniques [14]. Students' CT learning outcomes are assessed in the cognitive, technological, and ethical domains [18].

In collaboration with Malaysia Digital Economy Corporation (MDEC), the Ministry of Education (MOE) has designed CT TL-based professional development to train in-service teachers. Malaysian higher learning institutions are appointed to assist the implementation of the training programs. The training programs are implemented based on a training module organised and produced by MDEC. The training module provides an introduction and description of CT skills. The training program is implemented for at least a week, and participants are given follow-up tasks for an in-depth understanding of CT skills. Nonetheless, the training module does not emphasise on method or tool to assess CT TL outcomes.

2.2 CT Skills Assessment Method

CT researchers have conducted considerable research to pursue the best assessment method or tool based on the evidence of learning shown by learners. Most CT formal assessments focus on the middle school grades and above [19-21]. Progression of Early Computational Thinking Model (PECT) is proposed by Seiter and Foreman [22] in assessing CT learning outcomes of young children via computer programming. It focuses on the knowing state of the learners before and after a lesson is conducted. It is a model showing a learner's progression through the classes in enabling them to adjust their instructional strategy to cater to the teaching and learning.

Brennan and Resnick [23] suggest a few CT measurement tools and methods: students' portfolios, artefact-interviews, and design-scenarios. They propose a CT learning framework comprising three dimensions whereby CT TL outcomes being assessed based on computational concepts, computational practices, and computational perspectives. It must be noted that this work is carried out on programming lessons using Scratch as the TL platform. While the studies show promising results in assessing students' CT learning outcomes, the process is lengthy and tedious.

p-ISSN 1675-7939; e-ISSN 2289-4934

By using a different approach, Mark Dorling [24] implements Computing Progression Pathways, a rubric with 152 learning statements representing learners' performance and progress. This method matches the framework proposed by Curzon, et al. [25], but it may be arduous to assess a class with many students based on 152 learning outcomes.

Bers, et al. [26] introduce young children to robotic programming via Creative Hybrid Environment for Robotic Programming (CHERP). The learning outcomes are assessed based on a robot or program developed by a learner. The assessment is made based on CT concepts agreed on; learner achievement is scored based on a 6-point Likert scale. The achievement criteria are made based on the teachers' lesson outcomes. For instance, if a set of achievement statements tests students' proficiency in using control flow, another group of learning statements assess debugging while another set measures the repetitive flow. This process can be a challenge, as it will demand more time and effort spent to formulate the rubric each time when an assessment is carried out, which may be discouraging teachers from using it.

The literature review carried out indicates little work done on formal assessment methods that can be used to measure CT TL outcomes. The aforementioned methods are implemented in programming and computer-based lessons in a relatively small number of learners. There is no consensus on a uniform assessment method to be implemented, especially in the national curriculum. Tang, et al. [27] specifically express that more work should be carried out on the assessment method, especially for non-programming subject areas at all educational levels. It is vital to have a standardised and generic assessment tool that can be easily adapted into everyday classroom practices and applied to students in the Malaysian education environment.

3. PROBLEM STATEMENT AND OBJECTIVES

There is still no standard assessment tool available to assess students' performance in the CT concept, specifically for a rather large non-computing population such as primary school students in Malaysia. There is a need for a generalised, valid, and reliable tool to assess continuous CT TL outcomes. Besides, it has to be sufficiently robust to be implemented for assessing written tests and other means of assessment. Without a standardised assessment tool, it can contribute to the unreliable and invalid indication of students' performance, which may hinder TL from being delivered correctly. Hence, the main objective of the study is to design and develop an assessment tool to measure the continuous learning of CT across all taught subjects in primary schools.

A rubric, later known as myCTRubric, is proposed as an assessment tool due to its multi-dimensional set of scoring guidelines that enable teachers to give feedback and engage with the learners on their learning progress [28, 29]. It enriches communication among learners, schools, and parents. A simple one-page scoring rubric is suitable as it can help teachers to evaluate CT skills after TL activities. It aims to provide a transparent, objective, and user-friendly assessment process.

4. METHODOLOGY

The study adopts an action research design used by Susman and Evered [30] as a primary guide to structure the course based on the myCTRubric design and development. It is a well-established and appropriate approach related to educational practices to ensure successful

p-ISSN 1675-7939; e-ISSN 2289-4934

teachers' professional growth, understanding, and awareness, which are crucial in the learning process [31]. Figure 1 illustrates the flowchart for the process of developing myCTRubric.

Figure 1: Research flow chart.

The myCTRubric development is started with checking and reading any documents related to Malaysia's newly revised curriculum 2017. They comprise official curriculum documents from the Ministry of Education (MOE); for example, the *Kurikurlum Standard Sekolah Rendah KSSR (Semakan)* [32], research publications authored by [33-35], and teachers' training resources. To pursue in-depth information regarding CT integration across all subjects, interviews with a few subject matter experts (SMEs) were carried out. They are master trainers of CT and CT experts from the MOE and MDEC. The primary purposes of the interviews are:

- to seek SMEs' opinion on CT in the new curriculum;
- to make sense of CT concepts and approaches to assess ? based on the national curriculum; and
- to gauge the issues and potential solutions related to CT assessment.

A simple text analysis [36] was executed to process the review. Student learning outcomes of CT were explored and CT skill set was identified.

4.1 Identifying CT Concepts

Based on the Malaysian curricula and the MDEC training module, six CT concepts were identified. They consist of decomposition, pattern recognition, abstraction, algorithm, logical reasoning, and evaluation. Figure 2 shows the elements of the CT skill set defined by MOE.

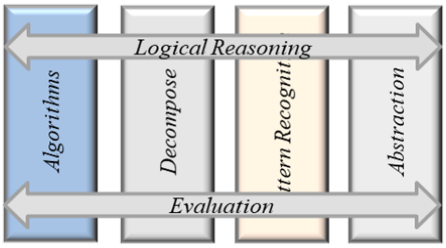

Figure 2: CT skill set listed by Ministry of Education Malaysia [37].

p-ISSN 1675-7939; e-ISSN 2289-4934

4.2 Mapping CT Elements Outcomes to National Learning Outcomes Basesd on CT Concepts

The newly revised curriculum documents (*Kurikulum Standard Sekolah Rendah KSSR (Semakan))* do not specify the CT learning outcomes. As such, primary school students performance reporting templates were referred to in the process of forming myCTRubric from the Standard Curriculum and Assessment Document (*Dokumen Standard Kurikulum dan Pentaksiran*) of Primary One. The six-level performances depicted in the Standard Curriculum and Assessment Document indicate the aggregate achievement of the pupils starting from Level 1 (Basic Users) to Level 6 (Proficient Users). While different subjects of Primary One (Malay Language, English Language, Science, Mathematics, Art, and Health Education) in Malaysia use different interpretations of performance that offer common grounds as depicted in Table 1.

Table 1 - Mapping of learning outcome similarities across all Primary 1 subjects.

Performance level	Description
1	Demonstrates a basic understanding of the relevant subject.
2	Demonstratesbasic understanding and can provide feedback on the subject.
3	Demonstrates basic understanding and can provide relevant feedback on the subject.
4	Demonstrates understanding and can provide correct feedback on the subject.
5	Demonstrates understanding and able to provide the correct feedback with explanations on the subject.
6	Demonstrates understanding, able to provide correct feedback, and provide explanations that can help in decision-making on the relevant subject.

All subjects also have included communication skills within the performance standards.

4.3 Formulation of myCTRubric

Figure 3 illustrates the progressions in formulating myCTRubric. In the formation of myCTRubric, the six-level performance indicators were used to evaluate student CT learning outcomes compliant with the Standard Curriculum and Assessment Document. The performance indicator was defined to obtain the objective measurable performance applicable for various assessment modes. The indicator level started with a minimal understanding before progressing to excellent knowledge of the skills. The performance standard description was stated within each indicator scale. The description was composed using distinct keywords found in the CT definition. The keywords were identified via a text network analysis [36] performed on a document containing all CT definitions.

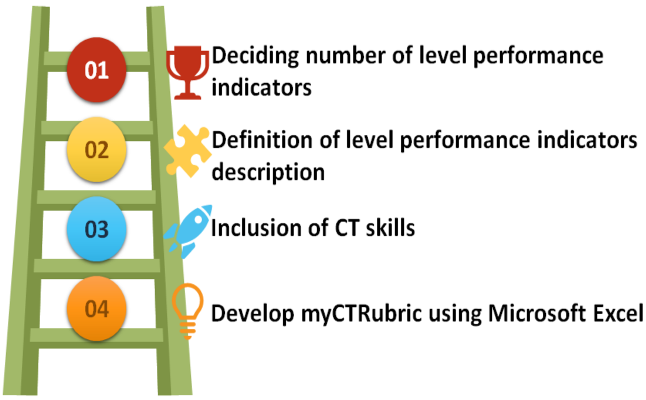

Figure 3: Formulation myCTRubric stages

p-ISSN 1675-7939; e-ISSN 2289-4934

The assessment of CT skills was based on six CT dimensions, which aredecomposition, pattern recognition, abstraction, algorithm design, logical reasoning, and evaluation. Students' performance in CT skills was measured on a scale of 1 to 6, which can be interpreted as 1: Very Limited, 2: Limited, 3: Fair, 4: Good, 5: Very Good, and 6: Excellent. Students who do not attempt finding solutions or do no execution, are rated as 1. It is not rated as '0' as SMEs believe that no execution might not signify a student does not understand the concept, but may not perform due to other issues as supported by the claims by Chen, et al. [38] and Childs, et al. [39].. Students are rated 2 when they solve a problem according to the teacher's instructions and rated 3 when they manage to solve a problem with minimal guidance. A student who is capable of executing a solution independently but exhibits some error is assigned as 4. To be rated as 5, a student can describe or justify their accomplished task. To be rated as 6, a student has to exhibit creator characteristics and produce a constructive conclusion.

myCTRubric was developed using Microsoft Excel 2010. The rows were labelled with the six CT skills and the columns were marked with a rating scale, starting with 6 as Excellent to the far right and 1 as Very Limited to the far left. Teachers can fill out the assessment results by marking on the relevant space. Space was provided at the bottom of myCTRubric for teachers to make notes on students' assessments. myCTRubric was first developed in English language version, and later the Malay language version was produced. The appendix shows the finalised myCTRubric.

4.4 Reviewing I

The first round of reviewing aims to ensure that myCTRubric can assess students' CT TL outcomes in line with the national learning goals. Three SMEs were invited (who were kept anonymous) to participate in the first reviewing stage. They reviewed, commented, and gave suggestions to improve the early version of myCTRubric.

4.5 Revisions

Iterations of feedback and revision were carried out. They evaluated the drafted myCTRubric mainly on language usage, clarity, and appropriateness to measure the content and CT skills [40, 41].

4.6 Reviewing II

At the second stage of the review, the finalised myCTRubric and an evaluation form were sent to another two SMEs (who were not involved in the first reviewing process). Both SMEs were asked to grade the two written tests, Test 1 and Test 2, on the Malay Language and Mathematics subjects using the finalised myCTRubric. The tests were answered by the same group of students. The result of the assessment then was used in the analysis to determine the rubric reliability. Intra-rater reliability was established using Intraclass Correlation Coefficients (ICC) to measure the degree of consistency and absolute agreement between measurements [42, 43]. Besides, the adjacent agreement, Cohen's Kappa, Spearman rank correlation, and Kendall rank correlation coefficients were calculated to measure the inter-rater agreement percentages between the two SMEs [44, 45]. A simple questionnaire was administered to gauge SMEs' perceptions of the proposed rubric. The questionnaire contained five Likert scale questions which were adapted from multiple sources [46, 47]. An open-ended section was provided to

p-ISSN 1675-7939; e-ISSN 2289-4934

allow the SMEs to state their concerns on the proposed rubric as a tool to assess their TL outcomes.

5. RESEARCH FINDINGS AND DISCUSSION

5.1 Reviewing I Finding

During the first reviewing process, most of the time was spent correcting grammatical mistakes and revising terms to ensure consistency in terms of keywords used in the description. For instance, “part” corrected to “parts”, and deciding the consistent application of keywords such as “identifies, breaks, simplifies”.

5.2 Reviewing II Findings

This section reports the outcome resulted from the second round of reviewing process. Two SMEs were involved in this process.

5.2.1 Intra-class Validity

A high degree of reliability is found between the two SMEs assessing both Test 1 and Test 2. The average measure of ICC for Test 1 is 0.968 with a 95% confidence interval from 0.869 to 0.961. While the average measure of ICC for Test 2 is 0.963 with a 95% confidence interval from 0.931 to 0.980.

Table 2 : Intraclass Correlation Coefficient for Test 1.

	Intraclass Correlation	95% Confidence Interval	
		Lower Bound	Upper Bound
Single Measures	0.928	0.869	0.961
Average Measures	0.962	0.930	0.980

Table 3 : Intraclass Correlation Coefficient for Test 2.

	Intraclass Correlation	95% Confidence Interval	
		Lower Bound	Upper Bound
Single Measures	0.929	0.871	0.962
Average Measures	0.963	0.931	0.980

5.2.2 Intra-rater Reliability

The adjacent agreement was computed to determine if there is an agreement between the two SMEs’ validation scoring. The SMEs’ scoring does not differ more than one point from the 6 performance standards. The adjacent agreement between the two SMEs is 83% for Test 1 and 88% for Test 2, indicating an acceptable adjacent agreement level [44, 48].

Cohen's kappa analysis indicates strong agreement between the two SMEs' assessment as K = 0.762 for Test 1 and K = 0.838 for Test 2. The Spearman rank correlation coefficients for Test 1 and Test 2 are 0.940 and 0.969, respectively. The Kendall rank correlation coefficients for Test 1 and Test 2 are 0.912 and 0.942, respectively. Both show excellent inter-rater reliability [49].

p-ISSN 1675-7939; e-ISSN 2289-4934

Table 4 : The results of Spearman Rank Correlation, Kendall Rank Correlation and Cohen's Kappa.

Spearman rho		Kendall rho		Cohen's Kappa	
Test 1	Test 2	Test 1	Test 2	Test 1	Test 2
0.940	0.969	0.912	0.942	0.762	0.838

5.2.3 SMEs Perceptions

Table 5 depicts the SMEs' perception of myCTRubric. Both have rated 5 (Strongly agree) in all statements except 4 (Agree) for the statement "The rubric can help me to plan for the next TL strategy." SME 1 thought that myCTRubric was relevant to the newly revised curriculum of 2017, while SME 2 noted that myCTRubric could assist in reporting students' CT learning outcomes.

Table 5 : SMEs' perceptions towards myCTRubric.

	SME 1	SME 2
I managed to measure students' CT TL outcomes.	5	5
I did not face any difficulties while assessing students' CT TL outcomes.	5	5
The rubric is suitable for measuring CT TL outcomes of this subject.	5	5
The rubric helps me to plan for the next TL strategy.	4	5
I will recommend other teachers to use this rubric in assessing CT TL outcomes.	5	5

6. DISCUSSION AND CONCLUSIONS

A generic rubric to evaluate CT skills is proposed across all subjects for Malaysian primary school students. We believe myCTRubric can provide valuable information, especially to teachers, parents, and education policymakers, on the effectiveness of CT TL in our national schools. This study offers an assessment instrument to quantify CT TL outcomes, which may aid in better CT TL design and self-review by teachers, schools, and even the MOE for better alignment with what is considered current standard best practice. The existence of myCTRubric is expected to contribute to a consistent and fair assessment. Also, this study will provide useful information for CT TL education, research, and instruction.

After reviewing myCTRubric, all the SMEs have confirmed its clarity, reliability, and validity. The validation tests show a statistically significant positive result. Even though adjacent agreement is often reported to produce positive results, many researchers prefer to use it in measuring inter-rater reliability because the agreement between raters is difficult to achieve [50, 51]. Cohen's Kappa analysis, Spearman rank correlation and Kendall rank correlation were carried out to reinforce the findings [52]. SMEs who participated in the reviewing process responded positively to implement the rubric in their assessment process.

However, there are a few limitations in this study. There are only a few SMEs involved in the reviewing process. The reliability and validity checks of myCTRubric can be improved by involving more SMEs as reviewers. Another limitation is the relatively limited number of assessment samples. The usability of myCTRubric can be further validated with more assessment samples in various assessment modes. An empirical investigation of myCTRubric is still in progress. myCTRubric will be piloted with a group of teachers in Malaysia. It is hoped

p-ISSN 1675-7939; e-ISSN 2289-4934

that this enquiry will provide school teachers with additional insights into more comprehensive CT TL strategies.

APPENDIX A: FINALISED MYCTRUBRIC

	EXCELLENT (6)	VERY GOOD (5)	GOOD (4)	FAIR (3)	LIMITED (2)	VERY LIMITED (1)
Decomposition	Breaks a problem into smaller parts correctly, accurately, confidently & describes it systematically.	Breaks a problem into smaller parts correctly.	Breaks a problem into smaller parts with minor errors.	Breaks a problem into smaller parts with significant errors.	Breaks a problem into smaller parts with teacher's guidance.	Displays no evidence of decomposition.
Pattern Recognition	Recognises patterns correctly and confidently, make accurate predictions based on the patterns/variations.	Recognises patterns correctly.	Recognises patterns with minor errors.	Recognises patterns with significant errors.	Recognises patterns with teacher's guidance.	Displays no evidence of pattern recognition.
Abstraction	Simplifies a problem, identifies the main ideas accurately, confidently, and describes the most important details of the problem.	Simplifies a problem and identifies the main ideas correctly.	Simplifies a problem and identifies the main ideas with minor errors.	Simplifies a problem and identifies the main ideas with significant errors.	Simplifies a problem and identifies the main ideas with teacher's guidance.	Displays no evidence of abstraction.
Algorithm Design	Creates a logical sequence of steps to solve a problem correctly, efficiently, confidently and able to describe the procedures in detail.	Creates a logical sequence of steps to solve a problem correctly.	Creates a logical sequence of steps to solve a problem with minor errors.	Creates a logical sequence of steps to solve a problem with significant errors.	Creates a logical sequence of steps to solve a problem with teacher's guidance.	Displays no evidence of algorithm.
Logical Reasoning	Predicts and explains the logical facts of a problem correctly, accurately and confidently.	Predicts logical facts of a problem correctly.	Predicts logical facts of a problem with minor errors.	Predicts logical facts of a problem with significant errors.	Predicts logical facts of a problem with teacher's guidance.	Displays no evidence of logical facts.
Evaluation	Evaluates the solution, identifies all mistakes/errors, successfully corrects and able to suggest improvements.	Evaluates the solution, identifies and corrects all mistakes/errors.	Evaluates the solution, identifies and successfully corrects most of the mistakes/errors.	Evaluates the solution, identifies and corrects some of the mistakes/errors.	Evaluates the solution with teacher's guidance.	Displays no evidence of evaluation.

REFERENCES

[1] D. Remy, H. Zahra, D. Katherine, S. Gerhard, and S. Philip, "Early informal STEM experiences and STEM identity: The importance of talking science," *Science Education*, vol. 103, no. 3, pp. 623-637, 2019.

[2] E. Ring-Whalen, Dare, E., Roehrig, G., Titu, P., & Crotty, E., "From Conception to Curricula: The Role of Science, Technology, Engineering, and Mathematics in Integrated STEM Units," *International Journal of Education in Mathematics, Science and Technology*, vol. 6, no. 4, pp. 343-362, 2018.

[3] U. Ling-Ling, Saibin, T. C., Labadin, J., & Aziz, N. A. , "Preliminary investigation: Teachers' perception on computational thinking concepts," *Journal of Telecommunication, Electronic and Computer Engineering* vol. 9, no. 2-9, pp. 23-29. [Online]. Available: https://journal.utem.edu.my/index.php/jtec/article/view/2672/1781

p-ISSN 1675-7939; e-ISSN 2289-4934

[4] A. Strawhacker and M. U. Bers, "Promoting Positive Technological Development in a Kindergarten Makerspace: A Qualitative Case Study," *European Journal of STEM Education*, vol. 3, no. 3, pp. 9, 2018.

[5] R. Tytler, "STEM Education for the Twenty-First Century," *Integrated Approaches to STEM Education*, pp. 21-43, 2020.

[6] S. Z. Abidina, Ibrahimb, A., Mohd, A. E., Ghazalic, M. A. A., Baharie, S. A., Mujirf, M. S., & Amino, M. N. Z. M., "The Challenges of Developing a Validation Protocol for the New Curriculum of Making 4.0: Between Malaysian and European Education Systems," 2020.

[7] F. Ilenia and P. Claus, "End-User Software Engineering in K-12 by Leveraging Existing Curricular Activities," in *ICSOFT*, pp. 283-289, 2018.

[8] N. Ki Won, K. Hye Jeong, and L. Suyoun, "Connecting Plans to Action: The Effects of a Card-Coded Robotics Curriculum and Activities on Korean Kindergartners," *The Asia-Pacific Education Researcher*, vol. 28, no. 5, pp. 387-397, 2019.

[9] M. Inomova, "Methodology Of Using Game Technologies In Preschool Education," *Архив Научных Публикаций JSPI,* 2020.

[10] M. Román-González, J. Moreno-León, and G. Robles, "Combining Assessment Tools for a Comprehensive Evaluation of Computational Thinking Interventions," in *Computational Thinking Education*, S.-C. Kong and H. Abelson Eds. Singapore: Springer Singapore, pp. 79-98, 2019.

[11] A. Strawhacker and M. U. Bers, "What they learn when they learn coding: investigating cognitive domains and computer programming knowledge in young children," *Educational Technology Research and Development*, vol. 67, no. 3, pp. 541-575, 2019.

[12] H. Lateh and V. Muniandy, "ICT implementation among Malaysian schools: GIS, obstacles and opportunities," *Procedia - Social and Behavioral Sciences*, vol. 2, no. 2, pp. 2846-2850, 2010/01/01/ 2010, doi: 10.1016/j.sbspro.2010.03.426.

[13] Kementerian Pelajaran Malaysia, *978-967-0359-01-4, Dokumen Standard Dunia Sains dan Teknologi Tahun 1*, 2011.

[14] Bernama, "Pemikiran komputasional, sains komputer akan diajar di sekolah tahun depan," in *Utusan ONLINE*, ed. Putrajaya: Utusan ONLINE, 2016.

[15] Malaysia Digital Economy Corporation, *DUTA PENGGERAK DIGITAL*, K. P. Malaysia, ed. [Online].Available:https://www.mydigitalmaker.com/sites/default/files/programmes/Aktiviti%20Sokongan%20Duta%20Penggerak%20Digital%20170328.pdf. Accessed on: 23 Oct 2017.

[16] *KURIKULUM STANDARD SEKOLAH RENDAH:Bahasa Cina Sekolah Kebangsaan Dokumen Standard Kurikulum dan Pentaksiran Tahun 1*, 2016.

[17] A. Yadav, J. Good, J. Voogt, and P. Fisser, "Computational thinking as an emerging competence domain," in *Competence-based Vocational and Professional Education, Technical and Vocational Education and Training: Issues, Concerns and Prospects*, vol. 23, M. Mulder Ed. Dordrecht, the Netherlands: Springer, Cham, pp. 1051-1067, 2017.

[18] K. P. Malaysia. "Standard Kompetensi Digital (Digital Competency Standards - DCS)." http://btp.moe.gov.my/article/222 (accessed 16/7/2018), 2018.

[19] R. J. Haddad and Y. Kalaani, "Can computational thinking predict academic performance?," in *Integrated STEM Education Conference (ISEC), 2015 IEEE*, 7-7 March 2015, pp. 225-229, doi: 10.1109/ISECon.2015.7119929, 2015.

[20] S. Promraksa, Sangaroon, K., & Inprasitha, M., "Characteristics of computational thinking about the estimation of the students in mathematics classroom applying lesson study and open approach," (in English), *Journal of Education and Learning,* vol. 3, no. 3, pp. 56-66, 2014-09-122014.[Online].Available:http://search.proquest.com.ezaccess.library.uitm.edu.my/docview/1561498832?accountid =42518.

[21] M. Zapata-Cáceres, E. Martín-Barroso, and M. Román-González, "Computational thinking test for beginners: Design and content validation," in *2020 IEEE Global Engineering Education Conference (EDUCON)*: IEEE, pp. 1905-1914, 2020.

p-ISSN 1675-7939; e-ISSN 2289-4934

[22] L. Seiter and B. Foreman, "Modeling the learning progressions of computational thinking of primary grade students," presented at the Proceedings of the ninth annual international ACM conference on International computing education research, San Diego, San California, USA, 2013.

[23] K. Brennan and M. Resnick, "New frameworks for studying and assessing the development of computational thinking," in *Proceedings of the 2012 Annual Meeting of the American Educational Research Association*, vol. 1. Vancouver, Canada: American Educational Research Association, pp. 1-25, 2012.

[24] S. S. Mark Dorling. "Computing Progression Pathways Guidance." Magpie Education. https://magpie.education/computing_faq7b.html#10 (accessed 22 May 2019, 2019).

[25] P. Curzon, M. Dorling, T. Ng, C. Selby, and J. Woollard, "Developing computational thinking in the classroom: a framework," 2014.

[26] M. U. Bers, L. Flannery, E. R. Kazakoff, and A. Sullivan, "Computational thinking and tinkering: Exploration of an early childhood robotics curriculum," (in English), *Computers & Education,* vol. 72, pp. 145-157, 3// 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0360131513003059.

[27] X. Tang, Y. Yin, Q. Lin, R. Hadad, and X. Zhai, "Assessing computational thinking: A systematic review of empirical studies," *Computers & Education,* vol. 148, p. 103798, 2020/04/01/ 2020, doi: https://doi.org/10.1016/j.compedu.2019.103798.

[28] A. Kaplon-Schilis and I. Lyublinskaya, "Unpacking performance indicators in the TPACK levels rubric to examine differences in the levels of TPACK," in *Proceedings of Society for Information Technology & Teacher Education International Conference*, E. Langran and J. Borup Eds. Washington, D.C., United States: Association for the Advancement of Computing in Education, sec. Waynesville, NC USA, pp. 2074-2083, 2018.

[29] S. Papadakis, M. Kalogiannakis, and N. Zaranis, "Designing and creating an educational app rubric for preschool teachers," *Education and Information Technologies,* vol. 22, no. 6, pp. 3147-3165, 2017.

[30] G. I. Susman and R. D. Evered, "An assessment of the scientific merits of action research," *Administrative Science Quarterly,* vol. 23, no. 4 (Dec.1978), pp. 582-603, 1978.

[31] L. Norton, *Action research in teaching and learning: A practical guide to conducting pedagogical research in universities*. New York: Routledge, 2018.

[32] K. P. Malaysia. "Kurikurlum Standard Sekolah Rendah KSSR (Semakan)." http://jip-ipgktho.blogspot.com/p/kurikulum-standard-sekolah-rendah-kssr.html (accessed 10 Mac 2017.

[33] L. C. Soon and J. Mustafa, "Tahap pemikiran komputasional dan hubungannya dengan prestasi akademik pelajar PISMP di Sarawak," *Selangor Humaniora Review,* vol. 2, no. 1, pp. 33-43, 2018.

[34] N. H. Anuar, F. S. Mohamad, and J.-L. Minoi, "Contextualising Computational Thinking: A Case Study in Remote Rural Sarawak Borneo," *International Journal of Learning, Teaching and Educational Research,* vol. 19, no. 8, pp. 98-116, 2020.

[35] F. Zainudin, "Enhancing Students' Higher Order Thinking Skills (HOTS) via Computational Thinking and Digital Literacy," ed, 2016.

[36] D. Paranyushkin, "Identifying the pathways for meaning circulation using text network analysis." [Online]. Available: https://www.semanticscholar.org/paper/Identifying-the-Pathways for Meaning Circulation Paranyushkin/97ba9f29c55ccbdd963e01f98cf17e73998f0f7d?p2df

[37] Kementerian Pendidikan Malaysia, *Bahasa Cina Sekolah Jenis Kebangsaan (C) TAHUN 1 Dokumen Standard Kurikulum dan Pentaksiran*, 2016.

[38] T. Y. Chen *et al.*, "Metamorphic testing: A review of challenges and opportunities," *ACM Computing Surveys (CSUR),* vol. 51, no. 1, pp. 1-27, 2018.

[39] R. A. Childs, S. Elgie, A. Brijmohan, and J. Yang, "" What if I don't know the answer?" Fifth-grade Students' Responses to Uncertainty in Test-Taking," *Canadian Journal of Education*, vol. 42, no. 4, 2019.

p-ISSN 1675-7939; e-ISSN 2289-4934

[40] T. S. Hodges, K. L. Wright, S. A. Wind, S. D. Matthews, W. K. Zimmer, and E. McTigue, "Developing and examining validity evidence for the Writing Rubric to Inform Teacher Educators (WRITE)," *Assessing Writing,* vol. 40, pp. 1-13, 2019/04/01/ 2019, doi: https://doi.org/10.1016/j.asw.2019.03.001.
[41] C. J. Brame, "Rubrics: Tools to Make Grading More Fair and Efficient," in *Science Teaching Essentials*, pp. 175-184, 2019.
[42] J. Harris, N. Grandgenett, and M. Hofer, "Testing a TPACK-based technology integration assessment rubric," in *Society for Information Technology & Teacher Education International Conference*: Association for the Advancement of Computing in Education (AACE), pp. 3833-3840, 2010.
[43] T. K. Koo and M. Y. Li, "A guideline of selecting and reporting intraclass correlation coefficients for reliability research," *Journal of chiropractic medicine,* vol. 15, no. 2, pp. 155-163, 2016.
[44] M. Graham, A. Milanowski, and J. Miller, "Measuring and Promoting Inter-Rater Agreement of Teacher and Principal Performance Ratings," *Online Submission,* 2012.
[45] Y. M. Reddy and H. Andrade, "A review of rubric use in higher education," *Assessment & evaluation in higher education,* vol. 35, no. 4, pp. 435-448, 2010.
[46] T. Djambong and V. Freiman, "Task-Based assessment of students’ computational thinking skills developed through visual programming or tangible coding environments," in *13th International Conference on Cognition and Exploratory Learning in Digital Age (CELDA 2016)*, D. G. Sampson, J. M. Spector, D. Ifenthaler, and P. Isaias Eds. Mannheim, Germany: IADIS Press, pp. 41-51, 2016.
[47] A. Niosi, "Creating Rubrics for Assessment," 2012.
[48] R. Koslow, "Exact and adjacent inter-rater agreement associated with peer review of teaching," *Assessment Update,* vol. 30, no. 2, pp. 1-16, 2018, doi: 10.1002/au.30125.
[49] D. V. Cicchetti, "Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology," *Psychological Assessment,* vol. 6, no. 4, pp. 284-290, 1994, doi: 10.1037/1040-3590.6.4.284.
[50] A. Jonsson and G. Svingby, "The use of scoring rubrics: Reliability, validity and educational consequences," *Educational Research Review,* vol. 2, no. 2, pp. 130-144, 2007.
[51] V. K. Aldridge, T. M. Dovey, and A. Wade, "Assessing test-retest reliability of psychological measures," *European Psychologist,* 2017.
[52] A. A. J. van Helvoort, S. Brand-Gruwel, F. Huysmans, and E. Sjoer, "Reliability and validity test of a Scoring Rubric for Information Literacy," *Journal of Documentation,* vol. 73, pp. 305-316, 2017, doi: 10.1108/jd-05-2016-0066.

p-ISSN 1675-7939; e-ISSN 2289-4934