An-Najah University Journal for Research – B
Humanities
مجلة جامعة النجاح للأبحاث – ب
العلوم الإنسانية

Validity and Reliability of the Computational Thinking Scale in Mathematics Learning

Rino Richardo[1,*], **Ahmad Anis Abdullah**[1,2] **& Robiatul Adawiya**[3]

(Type: Full Article). Received: 30th Jan. 2025, Accepted: 28th Sep. 2025 Published: ××××. DOI: https://doi

Accepted Manuscript, In Press

Abstract: Objective: Computational thinking (CT) is a fundamental skill that is a prerequisite for many future professions. The development of CT attitude is essential, as it plays a crucial role and influences CT in the cognitive domain. However, measurement scales of CT attitude remain limited, particularly in mathematics learning. Therefore, this study aims to develop a Computational Thinking Scale–Mathematics Education (CTS-ME) instrument that can be used to measure attitudes toward CT. **Method:** This study employed a quantitative approach, involving 231 students as respondents. Content validity was examined using Aiken's V, while construct validity was assessed through Confirmatory Factor Analysis (CFA). Corrected item–total correlation (CICT), Cronbach's alpha, and McDonald's Omega were used to examine reliability. **Main Findings:** The results revealed that the Aiken's V value exceeded 0.80 for all items, placing them in the high category. The CFA results showed that the proposed model achieved adequate fit indices. The instrument also met reliability criteria based on the three reliability indicators. **Conclusions:** It can be concluded that this scale is a satisfactory tool for measuring students' attitudes toward computational thinking in solving problems in mathematics learning. The CTS-ME instrument is recommended for use in self-assessment and in mapping students' initial CT abilities to support the design of differentiated mathematics instruction. Future studies are advised to involve more diverse samples and to employ advanced analytical approaches, such as the Rasch Model or SEM, to further strengthen the instrument's validity and generalizability.

Keywords: Computational thinking, Mathematics Education, Scale, Validity reliability, CFA.

الصدق والثبات لمقياس التفكير الحاسوبي في تعلم الرياضيات

رينو ريتشاردو[1،*]**، وأحمد أنيس عبد الله**[2،1]**، وربيعة العدويه**[3]

تاريخ التسليم: (2025/1/30)، تاريخ القبول: (2025/9/28)، تاريخ النشر: ××××

الملخص: الهدف: يعد التفكير الحاسوبي مهارة أساسية تمثل متطلباً مسبقاً للعديد من المهن في المستقبل، ويعد تنمية اتجاه التفكير الحاسوبي أمراً ضرورياً لأنه يلعب دوراً محورياً ويؤثر في التفكير الحاسوبي في المجال المعرفي، ومع ذلك، فإن مقاييس قياس اتجاه التفكير الحاسوبي ما زالت محدودة، خصوصاً في تعلم الرياضيات. لذلك، تهدف هذه الدراسة الى تطوير أداة مقياس التفكير الحاسوبي التي يمكن استخدامها لقياس الاتجاه نحو التفكير الحاسوبي، **المنهج:** ولتحقيق هذا الهدف، استخدمت (CTS-ME) في تعليم الرياضيات بينما تم تقييم Aiken V (0.8) وألفا كرونباخ، وأوميغا ماكدونالد لاختبار الثبات. **أهم النتائج:** أظهر نتائج الدراسة أن قيمة (CICT) الكلى لجميع الفقرات، مما يضعها في الفئة العالية. كما بينت نتائج تحليل العوامل التوكيدي أن النموذج المقترح يتمتع بمؤشرات جودة ملائمة. كذلك، استوفى المقياس معايير الثبات وفقاً للاختبارات الثلاثة المستخدمة. **الاستنتاجات:** ويمكن الاستنتاج بأن هذا المقياس يعد أداة مرضية لقياس اتجاهات الطلبة نحو التفكير الحاسوبي عند حل المشكلات في تعلم الرياضيات. تُوصى أداة **CTS-ME** باستخدامها في التقييم الذاتي وفي تحديد القدرات الأولية للطلبة في مجال التفكير الحاسوبي، وذلك لدعم تصميم تعليم رياضيات متمايز .كما تُنصح الدراسات المستقبلية بتضمين عينات أكثر تنوعًا واستخدام نماذج تحليلية متقدمة مثل نموذج راش أو نمذجة المعادلات الهيكلية **(SEM)** لتعزيز صلاحية الأداة وقابليتها للتعميم.

الكلمات المفتاحية: التفكير الحاسوبي، تعليم الرياضيات، المقياس، الصدق والثبات، تحليل العوامل التوكيدي (CFA.

1 Mathematics Education Department, Universitas Alma Ata, Yogyakarta, Indonesia
* Corresponding author email: rinorichardo@almaata.ac.id
2 E-mail: ahmad.anis@almaata.ac.id
3 School of Education, University of Szeged, Hungary. Robiatul.Adawiya@stud.u-szeged.hu

1 قسم تعليم الرياضيات، جامعة ألما آتا، يوجياكارتا، إندونيسيا
* الباحث المراسل: المراسل: rinorichardo@almaata.ac.id
2 البريد الالكتروني: ahmad.anis@almaata.ac.id
3 كلية التربية، جامعة سيجد، المجر. Robiatul.Adawiya@stud.u-szeged.hu

Introduction

Present-day technological advancements require skills that are necessary for adapting to a computer-based learning environment, with computational thinking (CT) being a crucial one (Nouri *et al*., 2019). It is considered a basic skill that is required for various professions in the future (Kafai & Proctor, 2022). The reason for this is that computational thinking is able to develop several abilities, such as creativity, logical and critical thinking, as well as problem-solving of each individual (Saritepeci, 2020).

CT is not a computer-like way of thinking; rather, it involves the ability to formulate problems, compile solutions in the form of algorithms, identify certain patterns and representations to solve them, and figure out and explain any inappropriate solutions. CT is part of the problem-solving thinking process incorporating several stages: problem identification, data collection, solution development, solution implementation, and assessment (Rao & Bhagat, 2024). Skills in CT are closely related to mathematics education. Increasing students' ability to analyze, solve, and resolve problems in a variety of situations is the main goal of mathematics education (Yeni *et al*., 2024). Since 2022, Computational Thinking (CT) has been a key component of mathematical literacy in PISA assessments. CT requires logical, analytical, and reflective thinking skills to solve contextual problems.

Problem-solving is at the core of learning mathematics at school (Santos-Trigo, 2020; Al-Khateeb, 2013; Al-Balawneh, 2010). This process not only helps students master mathematical concepts, but also prepares them to face real-life challenges that require logical, critical, and creative thinking skills (Muchlis *et al*., 2023). In the context of mathematics learning, students often have difficulty applying computational thinking concepts to solve problems. Some of them find it difficult to analyze and break down word problems into systematic steps (Kite & Park, 2024), unable to recognize hidden patterns in a number of mathematical concepts (Ginat, 2012), difficult to simplify complex information into a more easily understandable representation (Rosali & Suryadi, 2021), design logical and sequential problem-solving procedures (Kuo & Hsu, 2020, and reflect on mistakes made and improve previously generated solutions (Nordby et.al., 2022).

The difficulties students face in applying CT lie in the cognitive domain, reflecting limitations in developing thinking processes. To improve this ability, it is not enough to only provide treatment in the cognitive domain, improving CT also needs to consider affective factors, such as students' attitudes towards CT, because attitude factors also influence the cognitive domain. This is in line with findings that the cognitive domain representing learning outcomes shows that attitudes influence and determine the quality of student learning outcomes (Smyth *et al*., 2020; Richardo *et al*., 2023a; Richardo *et al*., 2023b). Therefore, attitude is very important to be measured and developed through non-test instruments, such as the computational thinking scale (CTS).

By definition, the aspects of CT include decomposition, algorithmic thinking, abstraction, as well as pattern recognition and generalization (Lee at al., 2020; Ehsan *et al*., 2021). Abstraction is the ability to identify and select essential information from a given problem (Kafai & Proctor, 2022). Decomposition is the ability to deconstruct complicated issues to simplify them (Morze *et al*., 2021). Algorithmic thinking is the ability to organize solution steps based on predetermined sub-problems (Purwasih &

Dahlan, 2024). Meanwhile, pattern recognition and generalization are the ability to recognize the same or different patterns to build a general solution to a given problem (Ye *et al*., 2023).

Currently, CTS instruments have been developed and adapted by several researchers. The instrument is still limited to the general domain and specific to certain fields. However, there is no computational thinking scale available in the field of mathematics education. Previous researches related to CTS are presented in Table 1

Table (1): Development of the CTS Instrument Development from 2014 – 2024.

Study	Item	Indicators	Student of Participant	Specific domain
Korkmaz *et al* (2017)	29	Creativity, Algorithmic Thinking, Cooperativity, Critical Thinking, Problem Solving	Undergraduate students	General
Korkmaz and Bai (2019)	20	Creativity, Algorithmic Thinking, Cooperativity, Critical Thinking, Problem Solving	K-10 dan K-11	General
Sirakaya *et al*. (2020)	22	Creativity, Algorithmic Thinking, Cooperativity, Critical Thinking, Problem Solving	Secondary school students	General
Junpho *et al*. (2022)	22	Creativity, Algorithmic Thinking, Cooperativity, Critical Thinking, Problem Solving	Junior High School students	General
Gök and Karamete (2023)	19	abstraction, decomposition, algorithmic thinking, evaluation and generalization	Junior High School students	Computer science and mathematics
Kılıc *et al* (2021)	33	Evaluation, conceptual knowledge, algorithmic thinking	University students	Computer programming
Yıldırım and Uluyol (2023)	17	Abstraction, algorithm, automation, decomposition, Generalization/Evaluation	Primary school students	Computer
Kukul & Karataş (2019)	18	Reasoning, Abstraction, Decomposition and Generalization	Secondary school students	Programming
Tsai *et al*. (2021)	19	abstraction, decomposition, algorithmic thinking, evaluation and generalization	Junior high school students	Computer literacy education
Yağcı (2019)	42	Problem-solving, Cooperative Learning & Critical Thinking, Creative Thinking, and Algorithmic Thinking	Junior high school students	Robotic
Current study	15	abstraction, problem decomposition, algorithmic thinking, pattern recognition & generalization.	Junior high school students	Mathematics Education

Based on Table 1, there is no specific CTS instrument in mathematics education. However, CTS specifically designed for mathematics education remain very limited. Therefore, there is an urgent need to develop a CTS tailored to mathematics education, which represents the novelty of the current research. In this study, we adapted the CTS previously developed by Tsai *et al*. for computer education literacy (Tsai *et al*., 2021). The domain was then contextualized into mathematics education, specifically related to mathematical problem-solving. Therefore, this study aims to fill the gap as described earlier by testing the validity and reliability of the CTS for Mathematics Education (CTS-ME) instrument. To achieve this goal, two research questions were raised: (1) Does the CTS-ME fulfill the validity criteria? and (2) Does CTS-ME a reliable measure?

Method

Sample: This research employed a quantitative approach (Creswell & Creswell, 2017). The data were collected from two private junior high schools in Yogyakarta, Indonesia. A total of 231 students were selected as the samples in this study from a population of 340 students using a convenience sampling technique. The sample size has fulfilled the element of adequacy, being five times greater than the number of items (Lakens, 2022)

Instrument: The CTS-ME was developed by adapting the CTS for educational computer literacy proposed by Tsai *et al*., particularly for junior high school students, which has been proven valid and reliable. The CTS consists of 19 items with 5 factors: Evaluation, Abstraction, Decomposition, Generalization, and Algorithmic Thinking. The evaluation factor consists of 4 items, abstraction consists of 4 items, decomposition consists of 3 items, Generalization consists of 4 items and thinking algorithm comprises 4 items (Tsai *et al*., 2021).

In this study, only 4 factors were adopted, including Abstraction, Decomposition, Algorithmic Thinking, and Generalization. The Generalization factor was developed into Pattern Recognition and Generalization. This is because in mathematics learning, especially problem solving, the context of Generalization is obtained simultaneously when it has found certain patterns. Therefore, the factors in the CTS-ME development include Abstraction (AB), Decomposition (DC), Algorithmic Thinking (AT), and Pattern Recognition and Generalization (PG). CTS-ME is presented in the form of a Likert scale questionnaire with 5 response options, ranging from 1 (strongly disagree) to 5 (strongly agree). The statement items on the CTS-ME instrument are presented in Table 2

Table (2): Items of the CTS-ME instrument.

Factors and Statements
Abstraction
(A1) I picture mathematics problems from a broad perspective, rather than looking at the details
(A2) I picture the connection between different problems in learning mathematics
(A3) I try to figure out the focal points of a mathematics problem.
(A4) I try to analyze common patterns of different problems in mathematics learning
Decomposition
(D1) I think about finding solutions to problems in mathematics learning
(D2) I think about the structure of problems in mathematics learning
(D3) I think about how to divide a large mathematics problem into several smaller problems
Algorithmic Thinking
(B1) I look for a step-by-step procedure to solve a mathematics problem
(B2) I try to look for an effective solution to a mathematics problem.
(B3) I try to organize the steps for solving a mathematics problem.
(B4) I try to find a way to solve a mathematics problem.
Pattern Recognition and Generalization
(PG1) I tend to solve mathematics problems according to my experience.
(PG2) I try to use commonly used methods to solve different mathematics problems.
(PG3) I consider how to apply an existing solution to another mathematics problem.
(PG4) I try to apply familiar solutions to solve more problems in mathematics learning

Meanwhile, the CTS-ME scale was administered by gathering students as respondents in the classroom to participate in a group session. In this session, each respondent needed approximately 20 to 30 minutes to complete the scale. Data collection is carried out directly by distributing the CTS-ME scale questionnaire to respondents. In the context of research ethics in Indonesia, studies focusing on scale testing do not require formal ethical approval, but their implementation still adheres to basic ethical principles by obtaining official

permission from the relevant school authorities before the activity is carried out.

Data Analysis: After the data collection, the validity and reliability tests were carried out to ensure the developed instrument was up to the mark and of high quality (Tobón & Luna-Nemecio, 2021). There were several stages of validity testing implemented in this study. First, the content validity testing was carried out using the Aiken V index. This test ensures that each item on the instrument represents the concept or construct being measured. It was performed through the review process of experts who are competent in computational thinking, especially in the field of mathematics education. On the other hand, the content validity test seeks to determine whether the instrument items align with the theoretical definition and contribute to the formation of the expected factor structure. The assessment aspects included in the content validity test are content, construction, and language.

The construct validity and reliability tests were conducted through Confirmatory Factor Analysis (CFA) using Jeffreys's Amazing Statistics Program (JASP (Rogers, 2024). CTS is a multidimensional measurement tool consisting of four main factors, namely abstraction, algorithmic thinking, decomposition, and pattern recognition and generalisation. Therefore, CFA approach is highly suitable for testing the factor structure of the CTS. CFA is designed to evaluate constructs consisting of multiple dimensions, where each factor represents a distinct theoretical dimension and is measured through a set of indicator items associated with each latent variable. CFA is widely used and has become a standard procedure in validating multidimensional measurement scales (Brown, 2015; Kline, 2016; Bryant & Satorra, 2012; Fabrigar & Wegener, 2012).

The process began with conducting the Kaiser-Meyer-Olkin (KMO) test by evaluating the Measure of Sampling Adequacy (MSA) value. The KMO test was calculated partially from each factor. The test criterion stipulates that the MSA value must be greater than 0.5 for the factor analysis to proceed (Lorenzo-Seva & Ferrando, 2021). In this study, the MSA (Measure of Sampling Adequacy) value was used as the basis for assessing data adequacy prior to confirmatory factor analysis (CFA). Although outlier identification is one of the ideal steps in the pre-processing stage before the KMO test, this approach was not explicitly carried out. This is because the MSA value per item has shown adequate correlation between variables, so it is considered sufficient to assess the suitability of the factor structure. Methodologically, this approach is acceptable as long as there are no indications of significant disturbances in the correlation structure of the data (Kaiser, 1974; Hair *et al*., 2019; Tabachnick & Fidell, 2013).

Furthermore, the Bartlett test was performed, with the criterion that if the sig value <0.5, the correlation between factors or aspects within the sample is considered strong (DiStefano *et al*., 2019). The next step involved determining the loading factor value, aiming to assess the correlation between factors and their supporting variables (Saeed, 2022). The test criterion is that if the loading factor is greater than 0.4 (N = 321), the variables contribute significantly to the underlying factor (Xia & Yang, 2019). The cut-off values for loading factors based on the sample size used are presented in Table 3. These values are used as a reference for assessing the validity of indicators against the constructs measured. After that, the determination of the theoretical model of the designed CTS-ME was conducted to figure out whether the theoretical model is in accordance with the data. This can be assessed

through the fit index obtained and the level of fit received (Xia & Yang, 2019) presented in Table 4

Table (3): Cutoff Factor Loading Based on Sample Size.

Sample Size (N)	Cutoff
N ≥ 350	≥ 0.30 (Steven, 1992)
N ≥ 200	≥ 0.40 (Hair *et al*., 2019)
N ≥ 100	≥ 0.45 until 0.50 (Comrey & Lee, 1992)
N < 100	≥ 0.50 until 0.60 (Tabachnick & Fidell, 2019)

Table (4): The fit indices value.

Fit Indices	Fit Value Obtained	Accepted Fit
Root means square error of approximation (RMSEA)	0.08	≤ 0.08
Comparative Fit Index (CFI)	0.94	> 0.9
Bollen's Incremental Fit Index (IFI)	0.94	> 0.9
Relative Noncentrality Index (RNI)	0.94	> 0.9
Goodness of Fit Index (GFI)	0.95	> 0.9

Meanwhile, the reliability test was conducted using the corrected item-total correlation (CITC), Cronbach's alpha (CA), and McDonald's Omega (MO) values across four factors. For the CITC, a value above 0.3 indicates that the instrument meets the reliability criteria (Zijlmans *et al*., 2019). Likewise, for CA and MO, if the value is above 0.6, the instrument is considered reliable (Taber, 2018).

Result

Content Validity: The content aspect ensures that the instrument items align with the aspects of CT, including AG, DC, AT, and PG. The construction aspect evaluates whether the instrument has been well adapted and in accordance with the context of learning and mathematical problem-solving. Meanwhile, the language aspect assesses whether the language used in the instrument items follows proper rules and is written in clear, and easily understandable sentences.

Content validity involved five experts in the field of mathematics education, three of whom hold the title of professor and two of whom are associate professors. The results of the expert assessment were then analyzed using the Aiken index (Aiken, 1980). The analysis results indicate that all CTS-ME instrument items were valid with a high category. Detailed content validation results are shown in Table 5.

Table (5): Results of CTS-ME analysis with aiken index V.

Item / Indicator	R1	R2	R3	R4	R5	Aiken Index (V)	Category
A1	4	3	4	3	3	0.85	High
A2	4	3	4	3	3	0.85	High
A3	4	3	4	3	3	0.85	High
A4	4	4	3	3	3	0.85	High
D1	3	4	4	4	3	0.90	High
D2	4	4	4	4	3	0.95	High
D3	4	4	4	4	3	0.95	High
B1	4	3	3	4	3	0.85	High
B2	4	3	4	4	3	0.90	High
B3	4	3	4	4	3	0.90	High
B4	4	3	4	4	3	0.90	High
PG1	3	3	3	4	3	0.80	High
PG2	3	3	3	4	3	0.80	High
PG3	3	3	3	4	3	0.80	High
PG4	3	4	4	4	3	0.90	High

R_i = first to fifth validators

Table 5 demonstrates that the 15 items of the CTS-ME instrument have a high content validity value, implying that all experts agreed to give a high assessment of the item suitability to the concept being measured. The V value is an index of expert agreement on the item's alignment with the indicators to be measured. V value greater than 0.8 suggests that all items are relevant and effectively measure aspects in CT. Through the results of empirical data, it can be concluded that the CTS-ME demonstrates good content validity

Construct Validity: After confirming good content validity, the construct validity test was then carried out, involving 231 junior high school students. Factor analysis was employed for this test (Treviño, 2021). There were several steps undertaken, including the KMO test, Bartlett's test of sphericity, and factor loading analysis. The detailed results are presented in Table 6

Table (6): Result of the KMO test, Bartlett's test of Sphericity, and Factor Loading Analysis.

Item / indicator	Measurement of Sampling Adequacy (MSA)	P-value Bartlett's test of sphericity	Std. Est. (all) (Value of Factor Loading)
A1	0.663	< 0.001	0.355
A2	0.745		0.504
A3	0.881		0.613
A4	0.878		0.641
D1	0.819		0.628
D2	0.805		0.605
D3	0.892		0.524
B1	0.903		0.705
B2	0.873		0.688
B3	0.842		0.750
B4	0.873		0.597
PG1	0.810		0.460
PG2	0.802		0.475
PG3	0.839		0.630
PG4	0.832		0.693

Table 6 exhibits the analysis results, showing the KMO test for each indicator that constructs the CTS-ME. KMO is the initial requirement to determine whether factor analysis can proceed, based on the adequacy of the sample used. In other words, factor analysis can be continued if all indicators have an MSA value greater than 0.5. Based on the data analysis, it was found that all indicators had an MSA value exceeding 0.5, signifying that the factors or aspects within the samples are correlated. Additionally, sig value of Bartlett's test was recorded to be less than 0.5, indicating a high correlation.

The value of the loading factor determines the correlation between a factor and its supporting variables, showing whether these variables significantly contribute to the underlying factors. Each item's loading factor exceeded 0.4, except for item A1 under the AB factor, which has a value of 0.355. In other words, all items within the DC, AT, and PG factors contribute significantly, while only 3 items in the AB factor demonstrate significant contributions.

Based on these results, item A1, which does not meet the factor loading criteria, must be removed. Subsequently, the KMO test, Bartlett's test of sphericity, and factor loading were reanalyzed the detailed results are presented in Table 7.

Table (7): Result of KMO test, Bartlett's test of sphericity and Factor Loading after the removal of item A1.

Item/indicator	Measurement of Sampling Adequacy (MSA)	P-value Bartlett's test of sphericity	Std. Est. (all) (Value of Factor Loading)	Average Variance Extracted (AVE)	Composite Reliability (CR)
A2	0.845	< 0.001	0.456	0.400	0.881
A3	0.878		0.628		
A4	0.879		0.669		

Item/indicator	Measurement of Sampling Adequacy (MSA)	P-value Bartlett's test of sphericity	Std. Est. (all) (Value of Factor Loading)	Average Variance Extracted (AVE)	Composite Reliability (CR)
D1	0.810		0.635		
D2	0.779		0.596		
D3	0.890		0.526		
B1	0.903		0.706		
B2	0.871		0.687		
B3	0.848		0.752		
B4	0.878		0.598		
PG1	0.814		0.460		
PG2	0.801		0.478		
PG3	0.855		0.622		
PG4	0.859		0.698		

Based on Table 7, it was obtained that all items had MSA values above 0.5 and loading factor values above 0.4. The final step of the CFA test is to test the theoretical model of the CTS-ME. The output for the CTS model is presented in Table 8 and the model plot is illustrated in Figure 1

Table (8): The Fit Indices values of model.

Fit Indices	Obtained Fit Value	Accepted fit	Model
RMSEA	0.08	< 0.08	0.069
CFI	0.94	> 0.9	0.903
IFI	0.94	> 0.9	0.905
RNI	0.94	> 0.9	0.903
GFI	0.95	> 0.9	0.987

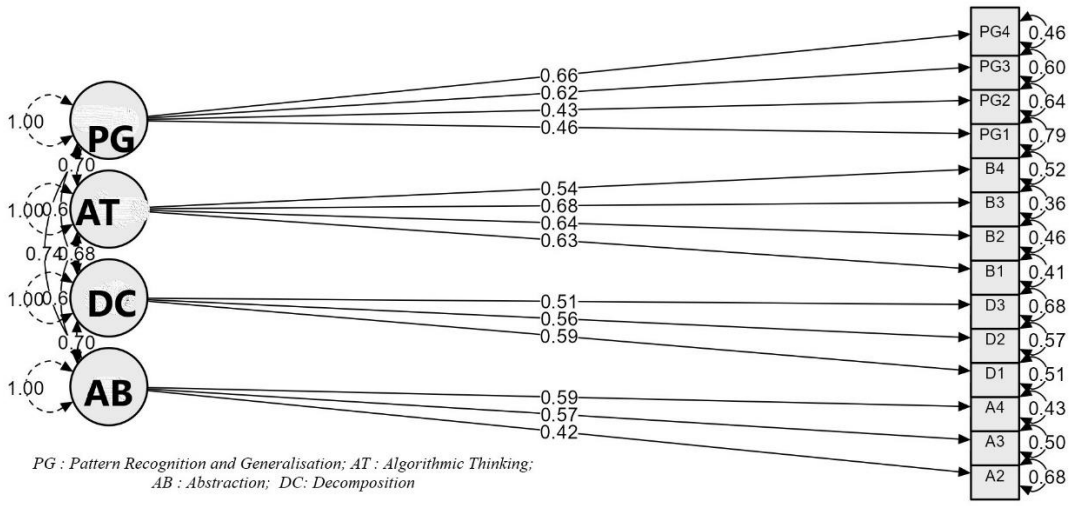

Figure (1): Plot of Model.

Based on Table 8, the fit indices indicators including RMSEA, CFI, IFI, RNI, and GFI have met the accepted fit value. In other words, the model satisfies the fit criteria. The CTS-ME construction in the model consists of four factors, comprising abstraction with 3 statement items, decomposition with 3 statement items, pattern recognition and generalization with 4 statement items, and thinking algorithm also with 4 statement items. In addition, the constructs in this study have a Composite Reliability (CR) value exceeding 0.80, namely 0.881, indicating a very high level of internal consistency between the indicators in measuring each construct. The

results of the measurement model analysis also show good fit, as reflected in various evaluation indices such as RMSEA, CFI, and GFI. These findings indicate that the instrument has adequate construct validity.

Reliability

The reliability test in this study employed three techniques including CITC, CA, and MO. These three techniques were chosen for their complementary nature, leading to a significant reliability value of the instrument developed. CITC evaluates each item on the scale, while, CA and MO assess the overall reliability. The results of the reliability test are presented in Table 9

Table (9): Result of Reliability Test.

Factor	Item / indicator	Corrected Item-Total Correlation (CITC)	McDonald's Omega (ω)	Cronbach's Alpha (α)
Abstraction	A2	0.396	0.608	0.603
	A3	0.491		
	A4	0.529		
Decomposition	D1	0.492	0.614	0.601
	D2	0.458		
	D3	0.445		
Algorithmic Thinking	B1	0.556	0.785	0.776
	B2	0.527		
	B3	0.569		
	B4	0.504		
Pattern Recognition and Generalization	PG1	0.394	0.653	0.649
	PG2	0.405		
	PG3	0.498		
	PG4	0.537		
Total			0.858	0.854

Based on Table 9, the CTS instrument was proven to be a reliable instrument. All items demonstrated a CITC value of more than 0.3 (Zijlmans *et al*., 2019). This means that all items in the CTS-ME exhibit a good correlation. Likewise, the CA and MO values of each factor exceeded 0.6, so the instrument is categorized as reliable (Taber, 2018).

Discussion

This study produced an instrument in the form of a scale called CTS-ME. This instrument can be used to measure students' CT attitude in learning mathematics at the junior high school level, particularly in the context of mathematical problem-solving. This instrument was developed by adapting the previously developed CTS. After the construct design of the CTS-ME was successfully developed, the instrument underwent content validation. This process involved a series of reviews to ensure whether all items were in accordance with the aspects measured by the indicators, as well as with the context and grammar used (Ermis-Demirtas, 2018). Content validation in this study was conducted by five expert validators in the field of mathematics education. Meanwhile, the content validity was analyzed using the Aiken index formula by looking at the level of agreement among expert assessments. The calculation results indicate that the content validity of each item exceeded 0.8, with an overall average of 0.87. This demonstrates that the instrument has

fulfilled the validity criteria in the high category (Angel-Cuervo, 2024).

Furthermore, the construct validity of CTS-ME, based on the CFA results, signifies that the suitability criteria for each item are within an adequate limit. The CITC value for each indicator was higher than 0.3, while the CA and MO values for each factor ranged from 0.601 to 0.785. The total CA value and MO value were recorded at 0.854 and 0.858, respectively. These values confirm that the CTS-ME satisfies the reliability criteria (Zijlmans *et al*., 2019; Taber, 2018). The results of confirmatory factor analysis indicate that the factor validity of the CTS-ME scale exhibited an adequate fit: RMSEA (0.069), CFI (0.903), IFI (0.905), RNI (0.903) and GFI (0.987). The CTS-ME model theoretically fulfils the fit index aspects with four aspects, namely abstraction, decomposition, algorithmic thinking, and pattern recognition and generalization.

These four aspects were designed using the context of mathematical problem solving at the junior high school level. These dimensions also form the basis for various CT scale instruments that have been developed previously by Korkmaz *et al*. (2017), Tsai *et al*. (2021), and Gök & Karamete (2023), demonstrating the consistency of the multidimensional construct of the CT scale across different educational levels. There are also similarities in terms of the respondents' educational level, namely junior high school students, which is also the focus of studies by Junpho *et al*. (2022), Tsai *et al*. (2021), and Yağcı (2019), indicating that this age group is important in the development of CT skills. However, unlike most previous studies that focused on the context of computer learning, programming, or digital technology, this study integrates the conditions of students' thinking processes in solving mathematical problems within the context of CT skills.

In general, the results of this study indicate that the developed CTS-ME instrument has met the criteria for validity and reliability, making it suitable for use as a non-cognitive measurement tool for self-assessment, measuring students' initial abilities in mathematics learning in the context of CT, thereby enabling educators to design differentiated teaching methods based on mastery of CT aspects and the development of positive attitudes towards mathematics. This study still has several limitations, including the fact that respondents were selected from private schools. Therefore, further research could address these limitations by including respondents from public schools, a combination of both public and private schools, and using a larger sample size. In addition, alternative analysis models, such as the Rasch Model or structural equation modeling (SEM), can be employed to more effectively and comprehensively assess the validity and reliability of the instrument.

Conclusions

Computational Thinking Scale - Mathematics Education (CTS-ME) is a non-cognitive instrument that empirically meets the aspects of content validity, construct validity, and reliability. The construction of this scale consists of four factors: abstraction with three indicators, and decomposition, algorithmic thinking, and pattern recognition and generalization, each with four indicators.

The content validity indicated by the Aiken V value exceeds 0.8 for all items, falling into the high category. The CFA results show that the established model has adequate fit indices. CTS-ME has also met the fit index indicators, based on the RMSEA, CFI, IFI, RNI, and GFI values, which have met the accepted fit values. The Composite Reliability (CR) value exceeds 0.80,

specifically 0.881, indicating a very high level of internal consistency among the indicators in measuring each construct. Regarding reliability, all items show a CITC value greater than 0.3, meaning all items in the CTS-ME demonstrate good correlation. Similarly, the CA and MO values for each factor exceed 0.6, so the CTS-ME is categorised as reliable.

Disclosure Statements

- **Ethical Approval and Consent to Participate:** Ethical approval for conducting this research was obtained from the participating school, and permission to collect data from students was granted by the school administration. All procedures were conducted in accordance with institutional ethical standards.
- **Data Availability:** The data supporting the findings of this study are available from the corresponding author upon reasonable request.
- **Authors' Contributions:** Author 1: Conceptualization, Methodology, Original Draft Writing, Supervision, Validation, Review & Editing. Author 2: Data Curation, Resources, and Investigation. Author 3: Editing, Visualization, and Formal Analysis.
- **Conflict of Interest:** The authors declare no conflict of interest.
- **Funding:** This research was funded by the Directorate of Research, Technology and Community Service (DRTPM) of the Ministry of Education, Culture, Research and Technology of the Republic of Indonesia in 2024, through the Fundamental Research Scheme, with Contract Number 107/E5/PG.02.00.PL/2024 and 016/A/SP3/LPPM/AA/VI/2024.
- **Acknowledgments:** The authors would like to express their gratitude to the Directorate of Research, Technology and Community Service (DRTPM), Ministry of Education, Culture, Research and Technology of the Republic of Indonesia, for providing funding support, and to the participating school for granting permission to conduct this research.

Open Access

References

- Aiken, L. R. (1980). Content validity and reliability of single items and questionnaires. *Educational and Psychological Measurement*, *40*(4), 955–959. https://doi.org/10.1177/001316448004000419
- Al-Balawneh, F. (2010). The effectiveness of an assessment strategy based on performance in developing mathematical thinking and problem-solving ability of secondary stage students. *An-Najah University Journal for Research – B (Humanities), 24*(8), 2227–

2270. https://doi.org/10.35552/0247-024-008-003

- Al-Khateeb, M. (2013). The effect of mathematical problem structure (context, content, number of problem-solving steps) on the ability to solve problems among second intermediate graders who have different mental capacities in Al-Madinah Al-Munawara. *An-Najah University Journal for Research – B (Humanities), 28*(7), 1629–1664. https://doi.org/10.35552/0247-028-007-005
- Angel-Cuervo, Z. M., Briceño-Martínez, J. J., & Bernal-Ballén, A. (2024). Validation of a questionnaire to evaluate mathematics teachers' beliefs about mathematics, teaching, and learning. *EURASIA Journal of Mathematics, Science and Technology Education, 20*(5), em2435. https://doi.org/10.29333/ejmste/14447
- Barahmand, A., & Attari, N. (2024). Investigating students' reasoning in generalization of deconstructive figural patterns. *Educational Studies in Mathematics,* 1-18. https://doi.org/10.1007/s10649-024-10355-z
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). New York, NY: The Guilford Press.
- Bryant, F. B., & Satorra, A. (2012). Principles and Practice of Scaled Difference Chi-Square Testing. *Structural Equation Modeling: A Multidisciplinary Journal*, *19*(3), 372–398. https://doi.org/10.1080/10705511.2012.687671
- Comrey, A. L., & Lee, H. B. (1992). *A first course in factor analysis.* (2nd ed.). Lawrence Erlbaum.
- Creswell, J. W., & Creswell, J. D. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches.* (5th ed.). Sage Publications.
- DiStefano, C., Zhu, M., & Mindrila, D. (2019). Understanding and using factor scores: Considerations for the applied researcher. *Practical Assessment, Research, and Evaluation, 14*(20).
- Ehsan, H., Rehmat, A. P., & Cardella, M. E. (2021). Computational thinking embedded in engineering design: Capturing computational thinking of children in an informal engineering design activity. *International Journal of Technology and Design Education, 31*(3), 441-464. https://doi.org/10.1007/s10798-020-09562-5
- Ermis-Demirtas, H. (2018). Establishing content-related validity evidence for assessments in counseling: Application of a sequential mixed-method approach. *International Journal for the Advancement of Counselling, 40*(4), 387-397. https://doi.org/10.1007/s10447-018-9332-4
- Fabrigar, L. R., & Wegener, D. T. (2012). *Exploratory factor analysis*. New York, NY: Oxford University Press.
- Ginat, D. (2012). Overlooking number patterns in algorithmic problem solving. In *Number Theory in Mathematics Education*. (pp. 223-247). Routledge.
- Gök, A., & Karamete, A. (2023). A validity and reliability study of the Turkish computational thinking scale. *Journal of Educational Technology & Online Learning, 6*(2), 421-437. https://doi.org/10.31681/jetol.1217363
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis.* (8th ed.). Andover, Hampshire: Cengage Learning.

- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis*. (8th ed.). Cengage Learning. https://doi.org/10.4324/9780429427507
- Junpho, M., Songsriwittaya, A., & Tep, P. (2022). Reliability and construct validity of computational thinking scale for junior high school students: Thai adaptation. *International Journal of Learning, Teaching and Educational Research, 21*(9), 154-173. https://doi.org/10.26803/ijlter.21.9.9
- Kafai, Y. B., & Proctor, C. (2022). A revaluation of computational thinking in K–12 education: Moving toward computational literacies. *Educational Researcher, 51*(2), 146-151. https://doi.org/10.3102/0013189X211057904
- Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika, 39*(1), 31–36. https://doi.org/10.1007/BF02291575
- Kılıç, S., Gökoğlu, S., & Öztürk, M. (2021). A valid and reliable scale for developing programming-oriented computational thinking. *Journal of Educational Computing Research, 59*(2), 257-286. https://doi.org/10.1177/0735633120964402
- Kite, V., & Park, S. (2024). Context matters: Secondary science teachers' integration of process-based, unplugged computational thinking into science curriculum. *Journal of Research in Science Teaching*, *61*(1), 203-227.
- Kline, R. B. (2016). *Principles and practice of structural equation modelling.* (4th ed.). New York, NY: The Guilford Press.
- Korkmaz, Ö., & Bai, X. (2019). Adapting computational thinking scale (CTS) for Chinese high school students and their thinking scale skills level. *Participatory Educational Research, 6*(1), 10–26. https://doi.org/10.17275/per.19.2.6.1
- Korkmaz, Ö., Çakir, R., & Özden, M. Y. (2017). A validity and reliability study of the computational thinking scales (CTS). *Computers in Human Behavior, 72,* 558–569. https://doi.org/10.1016/j.chb.2017.01.005
- Kukul, V., & Karataş, S. (2019). Computational thinking self-efficacy scale: Development, validity, and reliability. *Informatics in Education, 18*(1), 151–164. https://doi.org/10.15388/infedu.2019.07
- Kuo, W. C., & Hsu, T. C. (2020). Learning computational thinking without a computer: How computational participation happens in a computational thinking board game. *The Asia-Pacific Education Researcher*, *29*(1), 67-83.
- Lakens, D. (2022). Sample size justification. *Collabra: Psychology, 8*(1), 33267. https://doi.org/10.1525/collabra.33267
- Lee, I., Grover, S., Martin, F., Pillai, S., & Malyn-Smith, J. (2020). Computational thinking from a disciplinary perspective: Integrating computational thinking in K-12 science, technology, engineering, and mathematics education. *Journal of Science Education and Technology, 29*, 1–8. https://doi.org/10.1007/s10956-019-09803-w
- Lorenzo-Seva, U., & Ferrando, P. J. (2021). MSA: The forgotten index for identifying inappropriate items before computing exploratory item factor analysis. *Methodology, 17*(4), 296-306. https://doi.org/10.5964/meth.7185
- Morze, N., Boiko, M., & Barna, O. (2021). The relevance of training primary school teachers computational thinking. *ICT in Education, Research and Industrial Applications. Integration, Harmonization and*

Knowledge Transfer, *2*(3104), 141-153. https://ceur-ws.org/Vol-3104/paper218.pdf

– Muchlis, E. E., Priatna, N., & Maizora, S. (2023). Developing mathematical thinking skills through technology-based learning: A review of "technology-enabled mathematics education: Optimising student engagement." *Journal of Mathematics Teacher Education, 26*, 425–432. https://doi.org/10.1007/s10857-022-09561-4

– Nordby, S.K., Bjerke, A.H. & Mifsud, L. Computational Thinking in the Primary Mathematics Classroom: a Systematic Review. *Digit Exp Math Educ.* 8, 27–49 (2022). https://doi.org/10.1007/s40751-022-00102-5

– Nouri, J., Zhang, L., Mannila, L., & Norén, E. (2019). Development of computational thinking, digital competence and 21st century skills when learning programming in K-9. *Education Inquiry, 11*(1), 1–17. https://doi.org/10.1080/20004508.2019.1627844

– Purwasih, R., & Dahlan, J. A. (2024). How do you solve number pattern problems through mathematical semiotics analysis and computational thinking? *Journal on Mathematics Education, 15*(2), 403-430. https://doi.org/10.22342/jme.v15i2.pp403-430

– Rao, T. S. S., & Bhagat, K. K. (2024). Computational thinking for the digital age: A systematic review of tools, pedagogical strategies, and assessment practices. *Educational Technology Research and Development, 72*, 1893–1924. https://doi.org/10.1007/s11423-024-10364-y

– Richardo, R., Dwiningrum, S. I. A., & Wijaya, A. (2023a). Computational thinking skill for mathematics and attitudes based on gender: Comparative and relationship analysis. *Pegem Journal of Education and Instruction, 13*(2), 345-353. https://doi.org/10.47750/pegegog.13.02.38

– Richardo, R., Dwiningrum, S. I. A., Wijaya, A., Retnawati, H., Wahyudi, A., Sholihah, D. A., & Hidayah, K. N. (2023b). The impact of STEM attitudes and computational thinking on 21st-century via structural equation modelling. *International Journal of Evaluation and Research in Education, 12*(2), 571-578. https://doi.org/10.11591/ijere.v12i2.24232

– Rogers, P. (2024). Best practices for your confirmatory factor analysis: A JASP and lavaan tutorial. *Behavior Research Methods, 56*(7), 6634–6654. https://doi.org/10.3758/s13428-024-02375-7

– Rosali, D. F., & Suryadi, D. (2021). An analysis of students' computational thinking skills on the number patterns lesson during the covid-19 pandemic. *Formatif: jurnal ilmiah pendidikan MIPA*, *11*(2).

– Saeed, B., Tasmin, R., Mahmood, A., & Hafeez, A. (2022). Development of a multi-item operational excellence scale: Exploratory and confirmatory factor analysis. *The TQM Journal, 34*(3), 576-602. https://doi.org/10.1108/TQM-10-2020-0227

– Santos-Trigo, M. (2020). Problem-solving in mathematics education. In S. Lerman (Ed.), *Encyclopedia of Mathematics Education*. Springer, cham. https://doi.org/10.1007/978-3-030-15789-0_129.

– Saritepeci, M. (2020). Developing computational thinking skills of high school students: Design-based learning activities and programming tasks. *Asia-Pacific Education Researcher, 29*, 35–54. https://doi.org/10.1007/s40299-019-00480-2

- Sırakaya, M., Sırakaya, D. A., & Korkmaz, Ö. (2020). The impact of STEM attitude and thinking style on computational thinking determined via structural equation modeling. *Journal of Science Education and Technology, 29,* 561–572. https://doi.org/10.1007/s10956-020-09836-6
- Smyth, L., Mavor, K. I., & Platow, M. J. (2017). Learning behaviour and learning outcomes: The roles for social influence and field of study. *Social Educational Psychology, 20*, 69-95. https://doi.org/10.1007/s11218-016-9365-7
- Stevens, J. P. (1992). *Applied multivariate statistics for the social sciences*. (2nd ed.). Lawrence Erlbaum.
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics.* (6th ed.). Boston, MA: Pearson Education.
- Tabachnick, B. G., & Fidell, L. S. (2019). *Using multivariate statistics. (7th ed.).* Pearson. https://doi.org/10.4324/9781315814911
- Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education. 48*, 1273-1296. https://doi.org/10.1007/s11165-016-9602-2
- Tobón, S., & Luna-Nemecio, J. (2021). Complex thinking and sustainable social development: Validity and reliability of the COMPLEX-21 scale. *Sustainability. 13*(12), 6591. https://doi.org/10.3390/su13126591
- Treviño, M., Hu, X., Lu, Y. Y., Scheuer, L. S., Passell, E., Huang, G. C., ... & Horowitz, T. S. (2021). How do we measure attention? Using factor analysis to establish construct validity of neuropsychological tests. *Cognitive Research: Principles and Implications, 6*(51), 1-26. https://doi.org/10.1186/s41235-021-00313-1
- Tsai, M. J., Liang, J. C., & Hsu, C. Y. (2021). The computational thinking scale for computer literacy education. *Journal of Educational Computing Research. 59*(4), 579-602. https://doi.org/10.1177/0735633120972356
- Xia, Y., & Yang, Y. (2019). RMSEA, CFI, and TLI in structural equation modeling with ordered categorical data: The story they tell depends on the estimation methods. *Behavior Research Methods. 51*(1), 409-428. https://doi.org/10.3758/s13428-018-1055-2
- Yadav, A., Ocak, C., & Oliver, A. (2022). Computational thinking and metacognition. *TechTrends. 66*, 405–411. https://doi.org/10.1007/s11528-022-00695-z
- Yağcı, M. (2019). A valid and reliable tool for examining computational thinking skills. *Education and Information Technologies, 24*(1). 929-951. https://doi.org/10.1007/s10639-018-9801-8
- Ye, H., Liang, B., Oi-Lam, N., & Chai, C. S. (2023). Integration of computational thinking in K-12 mathematics education: A systematic review on CT-based mathematics instruction and student learning. *International Journal of STEM Education, 10*(3). https://doi.org/10.1186/s40594-023-00396-w
- Yeni, S., Grgurina, N., Saeli, M., Hermans, F., Tolboom, J., & Barendsen, E. (2024). Interdisciplinary Integration of Computational Thinking in K-12 Education: A Systematic Review. *Informatics in Education*. *23*(1), 223-278. https://doi.org/10.15388/infedu.2024.08
- Yıldırım, E., & Uluyol, C. (2023). Developing computational thinking scale for primary school students and examining students'

thinking levels according to different variables. *Journal of Learning and Teaching in Digital Age, 8*(1), 113–123. https://doi.org/10.53850/joltida.1176173

– Zijlmans, E. A. O., Tijmstra, J., Van der Ark, L. A., & Sijtsma, K. (2019). Item-score reliability as a selection tool in test construction. *Frontiers in Psychology, 9*, 2298. https://doi.org/10.3389/fpsyg.2018.02298.