AN ALGEBRAIC THINKING SKILL TEST IN PROBLEM-SOLVING FOR SEVENTH GRADERS

Janet Jahudin, Nyet Moi Siew
University Malaysia Sabah, Malaysia
E-mail: janetjahudin@gmail.com, sopiah@ums.edu.my

Abstract

Diagnostic tests have been developed previously to measure algebraic thinking skills; however, the tests do not specifically address algebraic problem-solving. Thus, an Algebraic Thinking Test (ATT) Instrument was developed to measure algebraic thinking skills in problem-solving involving linear equations. ATT comprises nine open-ended questions with three algebraic thinking constructs: Generalized Arithmetic, Functions and Modelling. Generalized arithmetic involves students in efficient calculation and generalization; functional involves identifying number patterns, while modelling involves solving open-ended problems, identifying similarities, and performing calculations involving variables. This study is meant to determine the quality of ATT instruments through the validity and reliability analysis using the Rasch Measurement Model. The sample consisted of 120 seventh graders aged 12 to 13, selected from two secondary schools in the Tuaran district, Malaysia. The instrument was found to have a strong dimensionality and high construct validity. The reliability of Cronbach Alpha (KR-20) demonstrated a value of .90 (very high), and item and respondent reliability of .98 (excellent) and .86 (good), respectively, with an item separation index of 6.29 and 2.45 for the person separation index. ATT has good validity and high reliability in measuring algebraic thinking skills among seventh graders in secondary schools.

Keywords: *algebraic thinking skills, linear equations, problem-solving, Rasch analysis, validity and reliability*

Introduction

Thinking skill is a distinct human capability. Nickerson (1985) defines thinking skills as a cognitive process an individual uses to solve problems. Beyer (1991) further elaborates that thinking skills are rooted in human abilities to form concepts, reason, and make decisions. Algebraic thinking skills are crucial components of mathematics education and have been the foundation of the subject for centuries (Carraher et al., 2007). Ibrahim and Othman (2010) suggested that algebra should be introduced to students at an early stage to stimulate their thinking abilities. For example, scholars such as Carraher and Schliemann (2007), Katz (2007), and National Council of Teachers of Mathematics (NCTM) (2000) argued that algebra, which involves reasoning, should be integrated into primary school curricular as a preparation for students to learn algebra in secondary school.

Mathematical processes are considered crucial for teaching and learning mathematics. These processes provide the groundwork for developing problem-solving abilities, which can be further enhanced through the application of various problem-solving strategies. As students learn to employ these strategies, they can develop critical thinking, creativity, and innovation skills.

In the past, the learning of algebra was mainly centred on symbol manipulation and arithmetic procedures, without much connection to the real world. According to Van Amerom

ISSN 1822-7864 (Print) ISSN 2538-7111 (Online) https://doi.org/10.33225/pec/23.81.223

(2002), this led students to perceive algebra as an abstract concept, making it challenging for them to understand. Additionally, the teacher focused on rote memorization of formulas to obtain correct exam answers without a deep understanding of the concepts and reasons. However, the current approach has shifted to training students in high-level thinking skills (HOTS) questions, which promote the development of conceptual, problem-solving, and application skills (Smith, 2014) related to algebraic thinking skills.

Research Problem

The average performance of Malaysian students in lower secondary schools in mastering algebraic concepts is still below par, as indicated by the results of the Trends in International Mathematics and Science Study (Mullis, 2019). Malaysia is ranked 29th out of 42 countries, and only a small percentage of students (5%) were able to apply algebraic problem-solving skills involving application, reasoning, and generalization. Similarly, studies in Finland, Sweden, and South Africa found that students struggled to master algebraic concepts, particularly equality and function. According to research by Bishop and Stump (2000), Jacobs et al. (2007) and Stephens (2008), one of the reasons for this difficulty is a lack of understanding of algebraic features. This deficiency is exacerbated by the fact that algebra is not formally introduced to students at the primary school level, which is emphasized more on arithmetic skills as reflected in Malaysian's Standard Primary School Curriculum.

In 2019, there has been a notable rise in emphasis on the development of algebraic thinking abilities in learning (Sibgatullin et al., 2022). Despite the attention given to this topic by researchers, there is a lack of instruments available to measure the level of algebraic thinking skills aligned with the Malaysian curriculum. The scarcity of assessment tools available to teachers for evaluating their students' algebraic thinking skills has contributed to this problem. Although a diagnostic test has been developed previously to measure generalization, function, and modelling skills (Ralston, 2013), it does not specifically address algebraic problem-solving. Consequently, the researcher developed a new instrument and assessed its validity and reliability to ensure its applicability in field research, hence to provide a reference for other researchers in the field to study the algebraic thinking abilities of secondary school students.

Literature Review

Algebraic Thinking Skills

Algebra is a branch of mathematics that employs symbols and letters as representations of values and quantities, and it plays a crucial role in other mathematical disciplines such as arithmetic, geometry, statistics, and calculus (Sianturi & Yang, 2021; Saundarajan et al., 2020). However, when students are formally introduced to algebraic topics at a lower secondary level, they often feel disinterested and not concerned (Lim & Lau, 2015). This leads to a barrier to developing algebraic thinking skills. Algebraic thinking skills refer to a process in learning mathematics that helps students to generalize mathematical ideas from examples to writing and arguments and statements (Hadi & Faradillah, 2019). Kaput et al. (2008) classified algebraic thinking skills into three main components: generalized arithmetic, function, and modelling. Ralston (2013) has further elaborated on the related dimensions of these three constructs. The construct of generalized arithmetic involves efficient calculation and generalization. The construct of functional thinking skills involves identifying number patterns, while modelling involves solving open problems, identifying similarities, and performing calculations involving variables.

Generalized Arithmetic

According to Kaput et al. (2008), generalized arithmetic is the study of algebra involving the examination of structures and systems that relate to computational and relational laws within mathematics. Fujii and Stephens (2001) suggested that students should be exposed to generalized arithmetic as an evolution from arithmetic to algebraic thinking skills. The development of algebraic thinking is described in some literature as a process that involves identifying and representing structures in mathematical expressions (Mason, 1989; Sfard & Linchevski, 1994). This process is related to an individual's cognitive ability to extract and establish generalizations, and to use appropriate forms of representation to express those generalizations (Carraher & Schliemann, 2007). This ability is the foundation for developing algebraic thinking skills, beginning with the basics of arithmetic (Carpenter et al., 2003; Jacobs et al., 2007).

Function

In Blanton and Kaput (2005)'s research, it was discovered that even primary school children have the ability to use different representations concerning function reasoning. This includes their capacity to verbally and symbolically describe recursive relations, covariations, and data matching and their ability to model and solve equations with unknown quantities using symbolic language. Blanton and Kaput's findings are supported by studies conducted by Brizuela and Lara-Roth (2000), Carraher et al. (2007), Moss and London McNab (2011), and Carraher and Schliemann (2007). In Lins and Kaput (2004)'s study with fifth-grade students, it was found that functional algebraic thinking skills can be developed through scaffolding. The gradual introduction of tables, graphs, pictures, words and symbols can shape students' thinking so that they can understand data and interpret functional relations more effectively.

Modelling

According to Ralston (2013), modelling is an essential component of algebraic thinking skills and involves solving open-number sentences, understanding similarities, and performing calculations involving variables. The use of variables and unknown values is indirectly introduced to students through activities that require them to find a value in an empty space, allowing them to understand the relations between arithmetic operations. Many researchers also emphasize equality as a means to develop algebraic thinking skills among students (Carpenter et al., 2005; Rittle-Johnson & Alibali, 1999). The symbol "equal to" has three distinct meanings: indicating that two quantities are equal, representing a relation, and suggesting the existence of two sides of an equation (Rittle-Johnson & Alibali, 1999, p.177). However, studies have shown that many students struggle with the conceptual understanding of variables, especially regarding the symbolic use of variables in algebra (Küchemann, 1978; Usiskin, 1988).

Algebraic Problem Solving

Solving algebraic problems is important because it helps students understand real-world situations and build meaning in real life (Kim & Chang, 2010; NCTM, 2000). However, students may require guidance in solving algebraic problems (Capraro & Joffrion, 2006), as these problems tend to be less popular among students due to their abstract nature (Baysal & Sevinc, 2021). Consequently, students often struggle with the understanding of problem statements, identifying variables, expressing known and unknown relations, comprehending information, and translating algebraic problems into expressions or equations, leading to a tendency to make mistakes (Clements, 1980; Egodawatte, 2011; Nathan & Koedinger, 2000).

Ralston et al. (2018) developed a diagnostic assessment tool with 25 items, containing eight components of algebraic skills related to 1745 elementary school students, demonstrating good item fit but low reliability. There is still a need for a comprehensive instrument that incorporates the constructs of generalized arithmetic, function, and modelling through problem-solving. To fill this gap, this study developed and tested an assessment tool for its validity and reliability that focuses on measuring these three constructs of algebraic thinking skills among high school students in terms of problem-solving ability.

Rasch Measurement Model

The Rasch Measurement Model (RMM) has been employed to measure validity and reliability in ATT items. This model is a useful tool for obtaining high instrument validity and reliability levels through advanced statistical techniques (Bond & Fox, 2007). The RMM assesses respondents' abilities to answer the instrument and evaluates the difficulty level of each item (Wolins, 1982). Additionally, the RMM can measure underlying characteristics such as human thoughts and emotions (Azrilah et al., 2015). While Rasch analysis may require a longer process than traditional analysis, it offers a more in-depth understanding of the strengths and weaknesses of the instrument (Boone & Scantlebury, 2006). According to Bond and Fox (2007), the RMM effectively creates highly valid and reliable instruments through statistical analysis.

The Rasch Measurement Model (RMM) is based on Item Response Theory and is considered an appropriate statistical model as it can simultaneously measure the difficulty and ability of individuals to respond to the items (Deane et al., 2016). The RMM analysis provides information about the reliability of individuals and items, the separation of item and person indices, and Cronbach's alpha value. Moreover, the construct validity of an instrument can be evaluated through item and person fit, as well as unidimensionality. The researcher used Rasch analysis to establish the reliability and validity of the ATT instrument by applying these key concepts.

Research Aim and Questions

The aim of this study was to develop a reliable and valid tool for assessing algebraic-related problem-solving abilities among seventh-grade students, using the Rasch Measurement Model. The study was guided by three research questions:

Q1: Is ATT feasible in assessing algebraic thinking skills related to problem-solving among seventh graders?

Q2: How valid is ATT in assessing seventh graders' algebraic thinking skills related to problem-solving?

Q3: How reliable is ATT in assessing the algebraic thinking skills involving problem-solving among seventh graders?

Research Methodology

Research Design

This research utilized a descriptive research design. The descriptive research design focused on evaluating the validity and reliability of the ATT instrument. The development of ATT involved a four-step process, which included construct identification, item formulation, pre-testing ATT, and validity and reliability testing. Prior to item construction, a literature review was conducted to gain an overview of the construct of algebraic thinking and the objectives of the national curriculum. This was followed by a comparison of the learning objectives of Mathematics in problem solving, as well as the results of TIMSS and PISA in algebra. Item pre-testing took place from 10 to 24 August 2022 to check if the questions and instructions were clear, after which modifications and refinements were made. The validity and reliability testing was conducted on 5 October 2022.

Sample

The research employed a purposive sampling technique. The selection of the research sample was based on the criteria that the students possessed a similar level of mathematics achievement as assessed by their teachers in the end-of-semester test marks (70-80%). The schools were categorized in the middle band in terms of mathematics achievement based on the Subject Average Grade set by the Tuaran District Education. A total of 120 seventh graders, consisting of 64 females (53.3%) and 56 males (46.7%) aged 12 to 13, were randomly selected from two rural secondary schools in the Tuaran district of Sabah, Malaysia. According to Linacre (1994), a sample size ranging from 108 to 243 can provide reliable and consistent data with a 99% confidence level for estimating item difficulty within ± ½ logit of its stable value.

Instrumentation

The Algebraic Thinking Test (ATT) consists of nine open-ended subjective questions. Multiple approaches can be taken to solve these subjective questions. As noted by Tofade et al., (2013), the process of answering these types of subjective questions is more vital to learning than attaining the correct answers as seen in objective questions. As such, these subjective questions promote high-level thinking skills. Table 1 gives more information regarding the ATT.

Table 1
ATT's Format

Format	Description
Skills tested	Algebraic problem solving based on the construct of Algebraic Thinking
Duration	1 hour 15 minutes
Number of Items	9
Language Use	Malay / English
Total Marks	45
Total Full Score (%)	100
Type of Item	Subjective (Open-ended response)
Cognitive level	Applying Analyze Evaluate

The ATT items were developed based on the content of the Curriculum Standard of Secondary School (CSSS) as well as the Standard Document of Curriculum and Assessment (SDCA) of Form One Mathematics under the field of Relations and Algebra (Curriculum Development Division, 2017: 53 – 56). Based on SDCA, with discussion and advice from the Coach of CSSS Mathematics and a Head of Mathematics, the researcher built a test specification table (TST) for one of the topics involved, as shown in Table 2.

Table 2
Algebraic Thinking Skills Test Specification Table

Item	Content Standards			Learning Standards				Performance Standard			Difficulty Level		
	6.1	6.2	6.3	6.1.2	6.1.3	6.2.2	6.32	GA	F	M	Low	Medium	High
1	/			/	/			/	/	/	/		
2		/				/	/	/	/	/		/	
3		/				/	/	/	/	/		/	
4	/			/	/			/	/	/	/		
5	/			/	/			/	/	/	/		
6	/			/	/			/	/	/		/	
7	/			/	/			/		/	/		
8		/				/	/	/	/	/			/
9		/				/	/	/	/	/			/
								Total			4	3	2

Each question in ATT includes three algebraic thinking skill constructs adapted from Kaput et al. (2008)'s algebraic thinking skills model: generalized arithmetic (GA), function (F), and modelling (M). Table 3 provides examples of questions based on each construct, while the solutions are illustrated in Figure 1. Each ATT item has a scoring system ranging from 0 to 6, with a minimum and maximum score of 0 and 6 respectively. The student's response determines the score awarded for each item, with a range of 0 to 6 marks for each item.

Table 3
The Constructs of ATT and Sample Questions

Sample Questions in ATT: The sum of three odd numbers in a row is 237. What is the number?		
Constructs	**Definition**	**Ability tested in the sample question**
Generalized Arithmetic (GA)	1. Using arithmetic laws and being able to manipulate numbers efficiently. 2. Doing generalized arithmetic.	Students can show an understanding that odd numbers have a sequence plus two from the previous odd number.
Function (F)	1. Specifying the relations between two or more variables in the form of quantity. 2. Expressing the relations between two or more variables as a number pattern.	Students can express the sequence of odd numbers in a pattern of numbers involving variables.
Modelling (M)	1. Using variables, especially in problem-solving. 2. Understanding the "equal" symbol means similarity.	Students can show linear equations

PROBLEMS
OF EDUCATION
IN THE 21st CENTURY
Vol. 81, No. 2, 2023

Figure 1
Demonstration of a Solution Using Conceptual Components of Algebraic Thinking

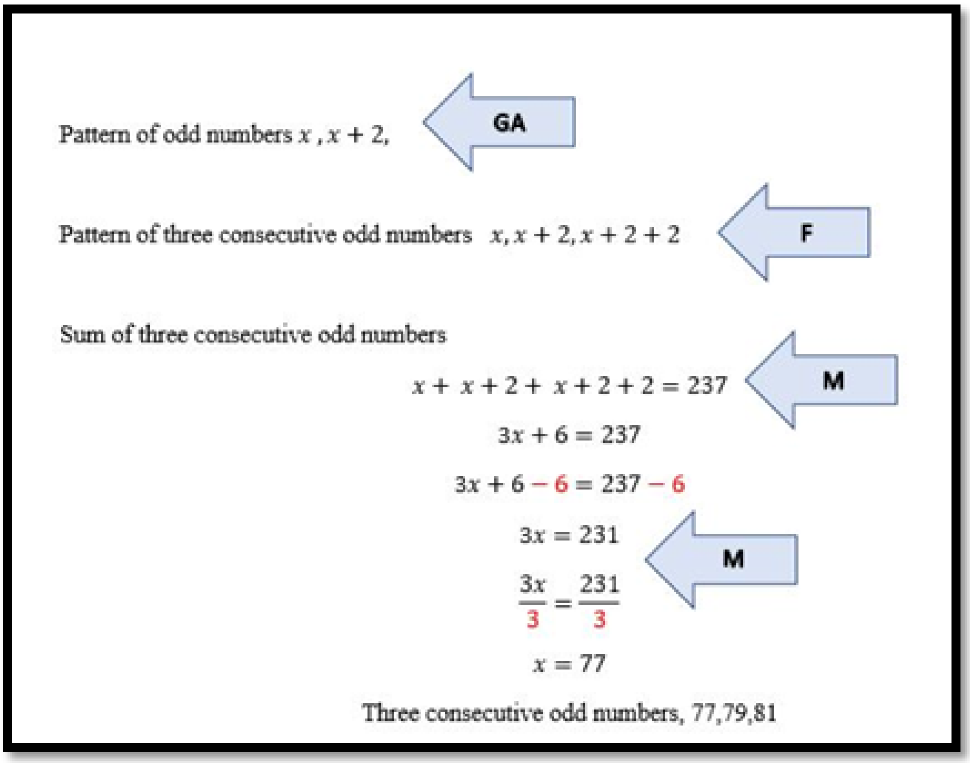
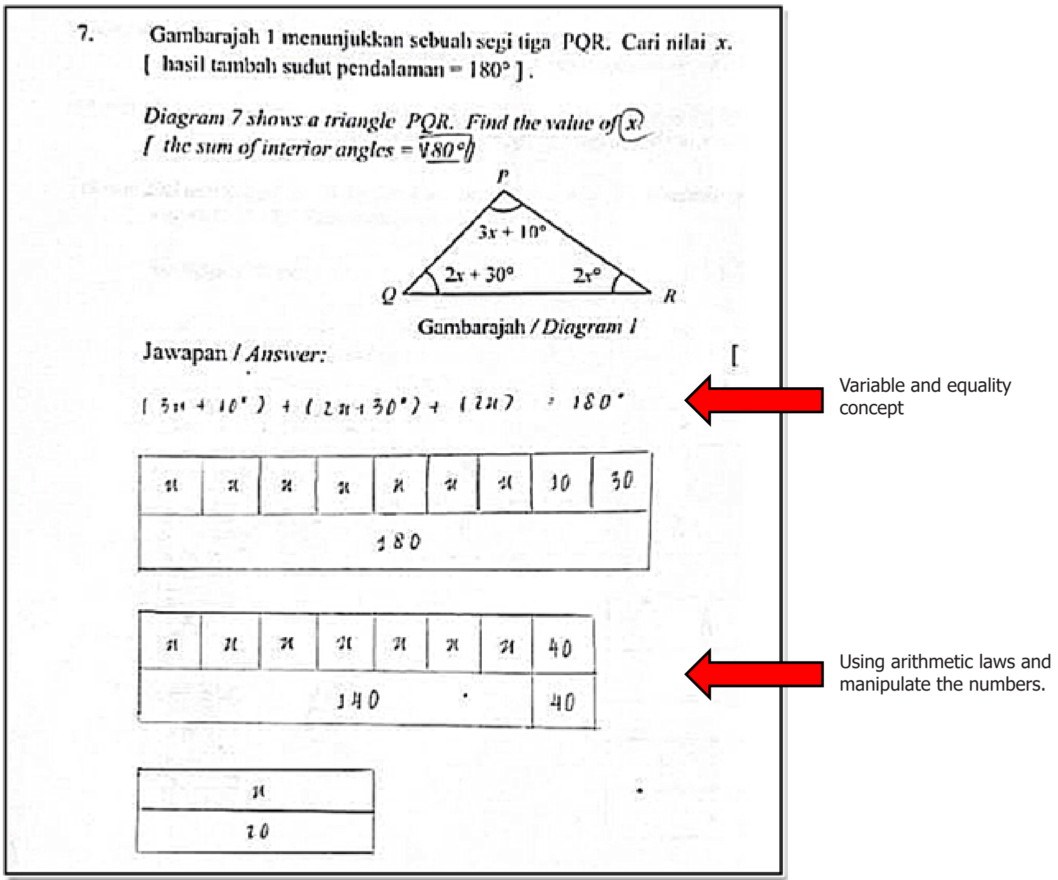


Figure 2 shows an example of the problem posed to the students in ATT. Attempting to solve the problem, the student was able to demonstrate the ability to apply the concept of variables in equality. The student also attempted to indicate the relation between variable and equality by showing that the sum of $7x$, 10, and 30 are the same as 180. To solve the problem, the student manipulated the equality concept and the numbers effectively by subtracting 40 from both sides of the equation, hence obtaining the value of x.

PROBLEMS
OF EDUCATION
IN THE 21st CENTURY
Vol. 81, No. 2, 2023

Figure 2
Student's Answer for Item 7

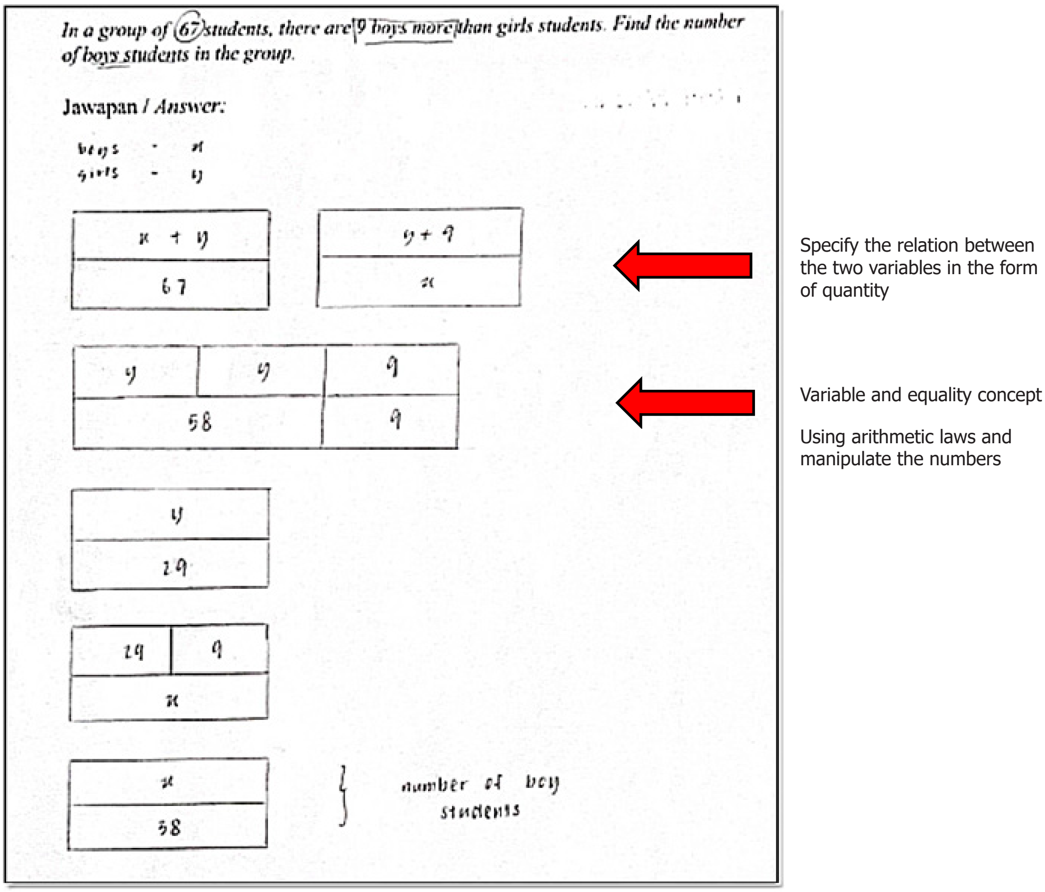


Meanwhile, in Figure 3, the student not only was able to demonstrate the ability to use variables and equality but also was able to specify the relation between the two variables.

https://doi.org/10.33225/pec/23.81.223 ISSN 1822-7864 (Print) ISSN 2538-7111 (Online)

PROBLEMS
OF EDUCATION
IN THE 21st CENTURY
Vol. 81, No. 2, 2023

Figure 3
Student's Answer for Item 8



Ethical Considerations

Prior to the administering of ATT, written consent was acquired from the participants' parents and school principals. The consent letter contained information about the students' involvement in the research, and the parents indicated that they had a clear understanding of the research objectives. Other than that, they were also provided with an outline of ATT's purpose to measure the level of algebraic thinking skills of the students. Participants were reassured that their responses would be kept confidential and anonymous, and that they were free to pull out of the study at any time without any consequences. Afterwards, a follow-up briefing was conducted to explain the instructions regarding how to answer the questions in ATT.

Administration of ATT

During the assessment, the students were instructed to read the ATT attentively, use their own knowledge to answer the questions, and avoid seeking assistance from external sources. They were given one hour to analyze and respond to the open-ended prompts. Following the completion of the task, the responses were collected, verified for correctness, and analyzed utilizing WINSTEPS software version 3.73.

ISSN 1822-7864 (Print) ISSN 2538-7111 (Online) https://doi.org/10.33225/pec/23.81.223

PROBLEMS
OF EDUCATION
IN THE 21st CENTURY
Vol. 81, No. 2, 2023

Content Validity

Validity pertains to the precision of the measurement in an instrument. Content validity evaluates the adequacy of the items in representing the trait's content that the researcher intends to measure, and experts' review ensures construct precision and content clarity, as cited by Creswell & Creswell (2017) and Kline (2005). To assess content validity value, the researcher utilized the Content Validity Index (CVI), which assesses the rating for all items evaluated by experts. Previous research specified that the Scale Content Validity Index (S-CVI) of .80 or higher (Davis, 1992) when the assessment of content validation involves three or more experts for a new instrument. This study applied a benchmark of .80, proposed by Davis (1992) for ATT. The analysis of CVI, based on Polit and Beck (2006), is presented below.

$$\text{Content Validity Index (CVI)} = \frac{\text{(Total score agreed by experts)}}{\text{(Total Score)}}$$

To enhance ATT's validity, the research conducted a content validity process with the assistance of four Mathematical Education experts. The researcher adopted an evaluation form suggested by the Malaysian Examinations Board to evaluate ATT items in terms of the suitability of algebraic thinking skills, language clarity, and item relevance. The experts' feedback was documented and used to improve the ATT items and refine their meaning, language, and presentation. Table 4 provides a list of the experts who participated in content validation.

Table 4
Content Validation Panel for ATT

Expert	Representative	Position	Expertise
Expert A	Public Universities	Associate Professor	Measurement and Assessment in Mathematical Education
Expert B	Public Universities	Academic Lecturer (PhD)	Mathematics Education, Statistics
Expert C	Institute of Teacher Education	Lecturer (PhD)	Mathematical Education
Expert D	Institute of Teacher Education	Lecturer Head Coach of KSSM Mathematics, Head Coach of Mathematics Educator, Mathematics' Examiner	HOTS in Mathematics, Mathematics Curriculum, Digital Mathematical Literacy and Mathematical Assessment

Table 5 shows that the Scale content validity index yielded a value of .94, surpassing the minimum requirement of .80, as Davis (1992) stipulated for new instruments. A S-CVI score of 0.94 indicates that the ATT instrument has a very high level of content validity, as assessed by the expert panel, and is deemed acceptable.

Table 5:
Content Validity Index (CVI) Result of the ATT instrument

Item	Expert				Assent Expert	I - CVI	Decision
	1	2	3	4			
1	/	/	/	/	4	1.00	Maintained
2	/	-	/	/	3	.75	Modified
3	/	-	/	/	3	.75	Modified
4	/	/	/	/	4	1.00	Maintained
5	/	/	/	/	4	1.00	Maintained
6	/	/	/	/	4	1.00	Maintained
7	/	/	/	/	4	1.00	Maintained
8	/	/	/	/	4	1.00	Maintained
9	/	/	/	/	4	1.00	Maintained
Scale Content Validity Index (S-CVI/ Average)						.94	

Construct Validity

The application of five statistical tests is required to conduct a basic Rasch analysis, which involves: (1) Measuring item polarity through Point Measure Correlation (2) Measuring item fit through the mean square (3) Confirming dimensionality (4) Determining Cronbach's alpha coefficient (KR-20) (5) Determining both item and person reliability by assessing their respective separation indices.

The item polarity is determined by the Point Measure Correlation (PTMEA-CORR) value analysis. A positive PTMEA-CORR value indicates that the item accurately measures its intended measure, and a negative value indicates the opposite. Additionally, item fit analysis was conducted using Outfit MNSQ, Outfit ZSTD, and PTMEA-CORR values, as suggested by Bond and Fox (2007), Boone et al. (2014), and Leitao and Waugh (2012). The purpose of item fit analysis is to determine if an item can perform its intended measurement (Sumintono & Widhiarso, 2015). Any items that fail to meet the criteria in Table 6 should be revised or removed to increase the item's fair value (Sumintono & Widhiarso, 2015).

To ensure the instrument's ability to measure algebraic thinking proficiency, the researcher conducted a dimensionality analysis using Principal Component Analysis (PCA), in accordance with Shea et al. (2009) and Sumintono and Widhiarso (2015). The dimensional criteria were based on the "raw variance explained by measures" value, which is considered good if it exceeds 20%, excellent if it exceeds 60%, and 40% or more is considered acceptable (Sumintono & Widhiarso, 2015). Additionally, the "unexplained variance in first contrast" value should not exceed 15%.

Table 6
Fit Indices for Item Fit

Statistics	Fit Indices
Outfit Means Square Values (MNSQ)	.50 – 1.50
Outfits Z-Standard Values (ZSTD)	-2.00 – 2.00
Point Measure Correlation (PTMEA CORR)	.40 – .85

Source: Boone et al. (2014)

Reliability

Sumintono and Widhiarso's (2015) work is referenced by the researcher for evaluating reliability in terms of Cronbach's Alpha (KR-20) values, as well as item-person reliability and separation indices, as shown in Table 7.

Table 7
Reliability Measured via Rasch Analysis

Statistics	Index	Interpretation
Alpha Cronbach (KR-20)	<.5	Low
	<.6	Moderate
	.6 – .7	Good
	.7 – .8	High
	.9 – 1.0	Very High
Item and Person Reliability Index	<.67	Low
	.67 – .80	Sufficient
	.81 – .90	Good
	.91 – .94	Very Good
	>.94	Excellent
Item and Person Separation Index	> 3.0	A high exile value indicates good persistence between the item and the person and describes the strata of the item and person in the instruments.

Source: Sumintono and Widhiarso (2015)

Research Results

Item Statistics

Linacre (2007) emphasized the importance of Outfit MNSQ values in evaluating the suitability of measurement items for a construct. For productive items, Boone et al. (2014) specified that the Outfit MNSQ range should be between .5 and 1.5, the Outfit ZSTD range should be within -2.0 to +2.0, and the PTMEA-CORR range should fall between .4 to .85. Based on Table 8, item 5's Outfit MNSQ value falls outside the range. As explained by Boone

https://doi.org/10.33225/pec/23.81.223 ISSN 1822-7864 (Print) ISSN 2538-7111 (Online)

et al. (2014) and Abdul Aziz et al. (2014), an item that does not meet all three criteria and falls outside the range is considered unsuitable. However, if an item meets any criteria, it should be retained, as suggested by Sumintono and Widhiarso (2015). Since all items meet at least one criterion in Table 8, no changes were made to the instrument.

Table 8
Item Statistics

Item	Outfit MNSQ (.50 - 1.50)	Outfit ZSTD (-2.0 - 2.0)	PTMEA-CORR (.40 - .85)	Decision
5	**1.94**	**5.10**	.46	Remain
1	.48	**- 4.70**	**.92**	Remain
6	.82	- 1.30	.72	Remain
7	.93	- .40	.81	Remain
2	1.10	.60	.68	Remain
3	.91	-.40	.67	Remain
4	.86	- .50	.66	Remain
8	.92	- .20	.68	Remain
9	.42	**- 2.10**	.75	Remain

Item Polarity

The assessment of item polarity was conducted using the PTMEA-CORR values, which determined whether the items in ATT moved congruently with the measured construct. A positive value indicated that all items functioned consistently, while a negative value indicated items that required improvement or elimination. As shown in Figure 4, all items had positive values, with PTMEA-CORR scores ranging from .46 to .92.

Figure 4
Analysis of Point Measure Correlation

```
INPUT: 109 Person  9 Item  REPORTED: 109 Person  9 Item  4 CATS    WINSTEPS 3.73
--------------------------------------------------------------------------------
Person: REAL SEP.: 2.45  REL.: .86 ... Item: REAL SEP.: 6.29  REL.: .98

        Item STATISTICS:  MEASURE ORDER

-------------------------------------------------------------------------------------------
|ENTRY   TOTAL  TOTAL           MODEL|   INFIT  |  OUTFIT  |PT-MEASURE |EXACT MATCH|      |
|NUMBER  SCORE  COUNT  MEASURE  S.E. |MNSQ  ZSTD|MNSQ  ZSTD|CORR.  EXP.| OBS%  EXP%| Item |
|------------------------------------+----------+----------+-----------+-----------+------|
|     5    370    109    71.41    1.75|1.85   4.9|1.94   5.1|  .46   .74| 43.1  66.3| I0005|
|     1    408    109    60.05    1.71| .51  -4.6| .48  -4.7|  .92   .74| 77.1  61.3| I0001|
|     6    413    109    58.60    1.71| .82  -1.4| .82  -1.3|  .72   .74| 62.4  60.0| I0006|
|     7    418    109    57.14    1.70| .99    .0| .93   -.4|  .81   .74| 47.7  60.3| I0007|
|     2    427    104    49.83    1.76|1.24   1.7|1.10    .6|  .68   .71| 64.4  64.1| I0002|
|     3    461    109    44.35    1.77|1.00    .1| .91   -.4|  .67   .70| 56.9  67.2| I0003|
|     4    473    109    40.48    1.82|1.03    .2| .86   -.5|  .66   .68| 69.7  67.5| I0004|
|     8    487    109    35.60    1.92| .88   -.8| .92   -.2|  .68   .65| 67.0  68.5| I0008|
|     9    495    109    32.54    2.00| .52  -3.9| .42  -2.1|  .75   .62| 91.7  71.1| I0009|
|------------------------------------+----------+----------+-----------+-----------+------|
| MEAN   439.1  108.4    50.00    1.79| .98   -.4| .93   -.4|           | 64.4  65.1|      |
| S.D.    39.5    1.6    12.11     .10| .38   2.7| .41   2.4|           | 13.9   3.7|      |
-------------------------------------------------------------------------------------------
```

Confirmation of Dimensionality

Determining the unidimensionality of an instrument is crucial to ensure accurate results are obtained from the study. The Principal Component Analysis of Rasch Residual (PCAR) is frequently used to detect unidimensionality. Based on the results presented in Figure 5, the Raw Variance Explained by Measures value's observed value is 61.1%, which according to Sumintono and Widhiarso (2015), is an excellent level of unidimensionality. This indicates that the ATT instrument is capable of measuring algebraic thinking skills. The unexplained variance in the first contrast was found to be 9.5%, which is below the threshold of 15%. Therefore, it can be concluded that the ATT instrument has strong unidimensionality.

Figure 5
Principal Component Analysis of Rasch Residual (PCAR)

```
INPUT: 109 Person  9 Item  REPORTED: 109 Person  9 Item  4 CATS    WINSTEPS 3.73
--------------------------------------------------------------------------------

    Table of STANDARDIZED RESIDUAL variance (in Eigenvalue units)
                                              -- Empirical --    Modeled
Total raw variance in observations     =       23.1 100.0%         100.0%
  Raw variance explained by measures   =       14.1  61.1%          60.8%
    Raw variance explained by persons  =        7.7  33.5%          33.3%
    Raw Variance explained by items    =        6.4  27.6%          27.5%
  Raw unexplained variance (total)     =        9.0  38.9% 100.0%   39.2%
    Unexplned variance in 1st contrast =        2.2   9.5%  24.5%
    Unexplned variance in 2nd contrast =        1.8   7.8%  20.2%
    Unexplned variance in 3rd contrast =        1.5   6.6%  17.0%
    Unexplned variance in 4th contrast =        1.2   5.3%  13.7%
    Unexplned variance in 5th contrast =         .9   4.0%  10.2%
```

Reliability and Separation Indices

To assess ATT's reliability, person measure reliability was used to distinguish students based on their responses, while item measure reliability was used to differentiate between items based on their approved authenticity by respondents. Table 9 presents an overview of the reliability findings. Based on Cronbach's alpha coefficient (KR-20), Table 9 indicates that ATT has strong reliability, with a value of .90. The item reliability score of .98 is classified as excellent (Sumintono & Widhiarso, 2015), and the item separation index of 6.29 is greater than 3.0, suggesting that ATT has a good distribution of items. According to Krishnan and Idris (2014), for instruments to accurately measure students, the separation of respondents should surpass 1.00. An item separation value above 2.00 is considered good (Linacre, 2003). Ardiyanti (2016) proposed an equation, H = [(4 x separation index) + 1] / 3, to determine the number of strata of separate items (H) based on the separation index. The resulting H value in this study is 8.72, rounded up to 9, indicating that ATT has nine levels of items.

PROBLEMS
OF EDUCATION
IN THE 21st CENTURY
Vol. 81, No. 2, 2023

Table 9
Summary of the Reliability Findings

Statistics	Value	Interpretation
Cronbach Alpha Rated (KR -20)	.90	High
Item Reliability	.98	Excellent
Item Separation Index	6.29	Good
Person Reliability	.86	Good
Person Separation Index	2.45	Good

In addition, Rasch's analysis also recorded a good respondent reliability value of .86 (Sumintono & Widhiarso, 2015). Bond and Fox (2007) also supported that the respondent's reliability rating higher than .80 indicates that respondents provide good and consistent feedback. The separation value of 2.45 is 'good' if its value exceeds 2.00 (Linacre, 2003). The separation value of 2.45 produces strata, H = 3.60 (can be rounded to 4). This value shows that students can be divided into four groups (weak, moderate, good and excellent) according to their responses based on items in ATT.

Figure 6
Principal Component Analysis of Rasch Residual (PCAR)

```
INPUT: 109 Person  9 Item  REPORTED: 109 Person  9 Item  4 CATS    WINSTEPS 3.73
--------------------------------------------------------------------------------

     SUMMARY OF 109 MEASURED Person
-------------------------------------------------------------------------------
|          TOTAL                         MODEL         INFIT        OUTFIT    |
|          SCORE     COUNT     MEASURE   ERROR      MNSQ   ZSTD   MNSQ   ZSTD |
|-----------------------------------------------------------------------------|
| MEAN      36.3       9.0       69.39    6.56       .96     .0    .93     .1 |
| S.D.       5.2        .2       18.90    1.53       .37     .8    .46     .8 |
| MAX.      44.0       9.0      103.90   10.97      1.62    1.3   2.10    1.3 |
| MIN.      25.0       8.0       27.47    5.58       .35   -1.7    .21   -1.7 |
|-----------------------------------------------------------------------------|
| REAL RMSE    7.14 TRUE SD   17.50  SEPARATION  2.45  Person RELIABILITY  .86 |
|MODEL RMSE    6.74 TRUE SD   17.66  SEPARATION  2.62  Person RELIABILITY  .87 |
| S.E. OF Person MEAN = 1.82                                                  |
-------------------------------------------------------------------------------
Person RAW SCORE-TO-MEASURE CORRELATION = .98
CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .90

     SUMMARY OF 9 MEASURED Item
-------------------------------------------------------------------------------
|          TOTAL                         MODEL         INFIT        OUTFIT    |
|          SCORE     COUNT     MEASURE   ERROR      MNSQ   ZSTD   MNSQ   ZSTD |
|-----------------------------------------------------------------------------|
| MEAN     439.1     108.4       50.00    1.79       .98    -.4    .93    -.4 |
| S.D.      39.5       1.6       12.11     .10       .38    2.7    .41    2.4 |
| MAX.     495.0     109.0       71.41    2.00      1.85    4.9   1.94    5.1 |
| MIN.     370.0     104.0       32.54    1.70       .51   -4.6    .42   -4.7 |
|-----------------------------------------------------------------------------|
| REAL RMSE    1.90 TRUE SD   11.96  SEPARATION  6.29  Item   RELIABILITY  .98 |
|MODEL RMSE    1.80 TRUE SD   11.97  SEPARATION  6.66  Item   RELIABILITY  .98 |
| S.E. OF Item MEAN = 4.28                                                    |
-------------------------------------------------------------------------------
```

ISSN 1822-7864 (Print) ISSN 2538-7111 (Online) https://doi.org/10.33225/pec/23.81.223

Discussion

The research conducted a comprehensive analysis of the validity and reliability of the ATT instrument for measuring algebraic thinking skills among seventh graders from secondary school, as summarized in Table 5 to Table 9. The validity analysis covered item suitability, polarity, and dimensionality.

After conducting an item conformity analysis, it was found that all items in the ATT instrument were deemed suitable for use as they met the criteria of Outfit MNSQ, Outfit ZSTD and PTMEA-CORR. A positive PTMEA-CORR analysis indicates that all instrument items are consistent in interpreting the measured constructs (Bond & Fox, 2007). Moreover, the Raw Variance Explained by Measures of the ATT instrument analysis shows a value of over 60%, indicating good dimensionality and confirming the instrument's ability to measure algebraic thinking skills in three constructs: generalized arithmetic, function, and modelling.

The reliability assessment of the ATT instrument indicates a high level of trustworthiness in measuring algebraic thinking skills among seventh graders. This is supported by the very high Cronbach's alpha value, the excellent reliability value of the instrument items, and the good reliability of the respondents. The ATT instrument also shows the good item and respondent separation values, indicating that the instrument has nine strata according to the number of items in ATT and the respondents are divided into four strata based on their level of ability (weak, medium, good, excellent).

While previous research has produced several instruments for assessing elementary students' level of algebraic thinking skills, such as those developed by Ralston et al. (2018), Alghtani et al. (2010), and Somasundram et al. (2016), there has been a need for a tool to assess these skills among secondary school students. ATT was developed to address this limitation and is explicitly designed to evaluate the level of algebraic thinking skills in problem-solving. Beyond its use in determining the extent of students' algebraic thinking skills, ATT also offers an important means of analyzing these skills in secondary Mathematics Education.

In general, the validity and reliability of ATT were evaluated using the Rasch Measurement Model. The ATT instrument focused on three constructs: Generalized Arithmetic, Functions, and Modelling. To ensure content validity, four experts reviewed all nine open-ended items. Item two was revised to avoid confusion in language writing. Additionally, the question's keyword was changed from "original perimeter of a garden" to "original perimeter of a rectangular shape garden" to encourage high-level thinking among students. The third item was modified to ensure that the meaning of sentences in Malay and English is consistent. Experts suggested that the question's intent should be aligned in both languages.

The research suggests that ATT, a new instrument in secondary mathematics education, aligns with Kaput's algebraic thinking skills model and can provide educators with valuable data on the extent of students' algebraic thinking skills. This data can be used to develop appropriate educational programs aimed at enhancing students' algebraic thinking abilities. Earlier studies, such as Hiebert et al. (1999), have established the significance of possessing algebraic thinking skills and have shown that it is a robust indicator of success in mathematics in the future. The findings of this study suggest that an appropriate algebraic thinking model can be used to assess the level of algebraic thinking skills among secondary students. Moreover, the ATT instrument utilizes problem-solving questions to enable learners to effectively tackle complex and non-routine mathematical problems. These results further support the idea that a strong foundation in algebraic thinking is critical for students to develop problem-solving skills and excel in mathematics, as emphasized by Fong and Lee (2009).

Furthermore, it is worth noting that while most previous instruments developed to assess algebraic thinking have used the exploratory factor analysis method, the ATT instrument employs

https://doi.org/10.33225/pec/23.81.223 ISSN 1822-7864 (Print) ISSN 2538-7111 (Online)

Rasch analysis. This approach is associated with greater accuracy in instrument development, which is an important advantage of ATT over previous instruments. Using Rasch analysis, the ATT instrument is better equipped to accurately measure students' algebraic thinking skills and provide valuable insights into areas where students may need additional support to improve their mathematical abilities.

Conclusion and Implications

Constructing an algebraic thinking skills instrument with a high level of validity and reliability is essential to ensure that the developed instrument can be used repeatedly. This is particularly important in enabling researchers to make informed decisions based on the instrument's findings. The results of this study also support the use of Rasch analysis in the development of the ATT instrument for measuring the algebraic thinking skills of students in other areas. By demonstrating the validity and reliability of the instrument using the Rasch Measurement Model, this study offers a strong foundation for its use in measuring the algebraic thinking skills of seventh-grade students in secondary schools.

This study has important practical and methodological implications. From a practical standpoint, the Algebraic Thinking Test (ATT) is a new instrument developed based on the algebraic thinking model. The algebraic thinking model can be adapted to suit the context of mathematics education in any country. The ATT instrument is well-suited to address the gap in mathematics education research, specifically in measuring algebraic thinking in problem-solving among secondary school students. The development of this instrument is a significant contribution towards ensuring the continuity of the application of algebraic thinking in mathematics education.

On the methodological side, the Rasch Measurement Model used to assess the validity and reliability of the ATT instrument is highly specific and detailed. This model's analysis indicates that the ATT instrument is highly valid for measuring algebraic thinking in problem-solving based on item-person fit, item polarity, and unidimensionality. High-reliability analysis of the ATT instrument, as indicated by a good item-person separation value, proves its reliability in measuring algebraic thinking involving three constructs: Generalized Arithmetic, Functions, and Modelling for seventh-grade students. Validity and reliability analysis of the ATT instrument is essential for accurate measurement of the variables studied and informed decision-making based on analysis of the findings. Furthermore, these findings provide a basis for other researchers to use the ATT instrument in assessing algebraic thinking in mathematics education among seventh-grade students in other contexts. Ultimately, the Rasch Measurement Model has successfully demonstrated the ATT instrument's suitability for actual field study in assessing algebraic thinking in mathematics education among seventh-grade students in secondary schools.

While the results demonstrate the reliability and validity of the ATT instrument for assessing algebraic thinking skills in secondary school, it is important to acknowledge its limitations. The study was conducted in rural schools and involved a sample size of only 120 students, which may not represent the broader population of secondary school students. Future research should expand the sample size to include urban schools to increase generalizability. Moreover, ATT can be integrated into STEM curriculum models to enhance critical and inventive thinking skills in Mathematics. To enhance the generalizability of the findings, it is recommended to extend the research of ATT to other regions and diverse student populations.

Acknowledgements

The researchers would like to express their appreciation to the Ministry of Higher Education Malaysia which funded this study under the Fundamental Research Grant Scheme (FRGS) Year 2021, FRGS/1/2021/SSI0/UMS/02/7.

Declaration of Interest

The authors declare no competing interest.

References

Abdul Aziz, A., Jusoh, M. S., Amlus, H., & Salleh, T. S. (2014). Construct validity: A Rasch measurement model approaches. *Journal of Applied Science and Agriculture*, *9*(12), 7-12.

Alghtani, O. A., & Abdulhamied, N. A. (2010). The effectiveness of geometric representative approach in developing algebraic thinking of fourth-grade students. *Procedia - Social and Behavioral Sciences*, *8*, 256–263. https://doi.org/10.1016/j.sbspro.2010.12.035

Ardiyanti, D. (2016). Aplikasi model Rasch pada pengembangan skala efikasi diri dalam pengambilan keputusan karier siswa [Application of Rasch's model of measurement for the self-efficacy scale in undergraduate's career decision making], *Jurnal Psikologi* [Psychological Journal], *43*(3), 248–263. https://doi.org/10.22146/jpsi.17801

Azrilah, A. A., Mohd Saidfudin, M., & Azami, Z. (2015). *Asas model pengukuran Rasch: pembentukan skala dan struktur pengukuran* [Rasch basic measurement model: scale formation and measurement structure]. Penerbit Universiti Kebangsaan Malaysia [UKM Press].

Baysal, E., & Sevinc, S. (2021). The role of the Singapore bar model in reducing students' errors on algebra word problems. *International Journal of Mathematical Education in Science and Technology*, *53*(2), 289–310. https://doi.org/10.1080/0020739x.2021.1944683

Beyer, B. K. (1991). *Teaching Thinking Skills*. Pearson.

Bishop, J. W., & Stump, S. L. (2000). Preparing to teach in the new millennium: Algebra through the eyes of pre-service elementary and middle school teachers. In M. Fernandez (Eds.), *Annual conference of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 107-113). ERIC Clearinghouse for Science, Mathematics, and Environmental Education.

Blanton, M. L., & Kaput, J. J. (2005). Characterizing a classroom practice that promotes algebraic reasoning. *Journal for Research in Mathematics Education*, *36*(5), 412–446. http://www.jstor.org/stable/30034944

Bond, T. G, & Fox, C. M. (2007). *Applying the Rasch model. Fundamental measurement in the Human science*. Lawrence Earlbaum Associates.

Boone, W. J., & Scantlebury, K. (2006). The role of Rasch analysis when conducting science education research utilizing multiple-choice tests. *Science Education*, *90*(2), 253–269. https://doi.org/10.1002/sce.20106

Boone, W. J., Staver, J. R., & Yale, M. S. (2014*). Rasch Analysis in the Human Sciences*. Springer.

Brizuela, B. M., & Lara-Roth, S. (2001). Additive relations and function tables. *The Journal of Mathematical Behavior*, *20*(3), 309–319. https://doi.org/10.1016/s0732-3123(02)00076-7

Capraro, M. M., & Joffrion, H. (2006). Algebraic equations: Can middle-school students meaningfully translate from words to mathematical symbols? *Reading Psychology*, *27*(2-3), 147–164. https://doi.org/10.1080/02702710600642467

Carpenter, T. P., Levi, L., Franke, M. L., & Zeringue, J. K. (2005). Algebra in elementary school: Developing relational thinking. *Zentralblatt Für Didaktik Der Mathematik*, *37*(1), 53–59. https://doi.org/10.1007/bf02655897

Carpenter, T. P., Franke, M. L., & Levi, L. (2003). *Thinking Mathematically*. Heinemann Educational Books.

Carraher, D. W., & Schliemann, A. D. (2007). Chapter 8: Is everyday mathematics truly relevant to mathematics education? *Journal for Research in Mathematics Education, Monograph*, *11*, 131-153. https://doi.org/10.2307/749968

Carraher, D. W., Martinez, M. V., & Schliemann, A. D. (2007). Early algebra and mathematical generalization. *ZDM*, *40*(1), 3–22. https://doi.org/10.1007/s11858-007-0067-7

Clements, M. A. (1980). Analyzing children's errors on written mathematical tasks. *Educational Studies In Mathematics*, *11*(1), 1-21.

Creswell, J. W., & Creswell, J. D. (2017). *Research design: Qualitative, quantitative, and mixed methods approach.* Sage.

Curriculum Development Division. (2017). *Mathematics Curriculum and Assessment Standards Document for Form One.* Malaysia Ministry of Education.

Davis, L. L. (1992). Instrument review: Getting the most from a panel of experts. *Applied Nursing Research*, *5*(4), 194–197. https://doi.org/10.1016/s0897-1897(05)80008-4

Deane, T., Nomme, K., Jeffery, E., Pollock, C., & Birol, G. (2016). Development of the statistical reasoning in biology concept inventory (SRBCI). *CBE—Life Sciences Education*, *15*(1), ar5. https://doi.org/10.1187/cbe.15-06-0131

Egodawatte, G., McDougall, D., & Stoilescu, D. (2011). The effects of teacher collaboration in Grade 9 Applied Mathematics. *Educational Research for Policy and Practice*, *10*(3), 189–209. https://doi.org/10.1007/s10671-011-9104-y

Fong, N. S., & Lee, K. (2009). Model method: A visual tool to support algebra word problem solving at the primary level. In K. Y. Wong, P. Y. Lee, B. Kaur, P. Y. Foong, S. F. Ng (Eds.), *Mathematics Education: The Singapore journey* (pp. 169–203). World Scientific. https://doi.org/10.1142/9789812833761_0008

Fujii, T., & Stephens, M. (2001). Fostering an understanding of algebraic generalization through numerical expressions: The role of quasi-variables. In H. Chick, K. Stacey, & J. Vincent (Eds), *Proceedings of the 12th ICMI study conference: The future of the teaching and learning of algebr*a (pp. 258–264). The University of Melbourne.

Hadi, W., & Faradillah, A. (2019). The algebraic thinking process in solving hots questions reviewed from student achievement motivation. *Al-Jabar Jurnal Pendidikan Matematika, 10*(2), 327-337. https://doi.org/10.24042/ajpm.v10i2.5331

Hiebert, J., Stigler, J. W., & Manaster, A. B. (1999). Mathematical features of lessons in the TIMSS video study. *ZDM*, *31*(6), 196-201.

Ibrahim, Z. B., & Othman, K. I. (2010). Comparative study of secondary mathematics curriculum between Malaysia and Singapore. *Procedia - Social and Behavioral Sciences*, *8*, 351–355. https://doi.org/10.1016/j.sbspro.2010.12.049

Jacobs, V. R., Franke, M. L., Carpenter, T. P., Levi, L., & Battey, D. (2007). Professional Development Focused on Children's Algebraic Reasoning in Elementary School. *Journal for Research in Mathematics Education*, *38*(3), 258–288. https://doi.org/10.2307/30034868

Kaput, J. J., Blanton, M. L., & Moreno, L. (2008). Algebra from a symbolization point of view. In Kaput, J. J., Carraher, D. W., Blanton, M. L. (Eds.), *Algebra in the early grades* (pp.19-56). Routledge.

Katz, V. J. (2007). *Algebra: Gateway to a technological future*. The Mathematical Association of America.

Kim, S., & Chang, M. (2010). Computer games for the math achievement of diverse students. *Journal of Educational Technology & Society*, *13*(3), 224-232.

Kline, T. J. (2005). *Psychological testing: A practical approach to design and evaluation*. Sage

Küchemann, D. (1978). Children’s understanding of numerical variables. *Mathematics in School*, *7*(4), 23–26. https://www.jstor.org/stable/30213397

Krishnan, S. & Idris, N. (2014). Investigating reliability and validity for the construct of inferential statistics. *International Journal of Learning, Teaching and Educational Research*, *4*(1), 51–60. https://doi.org/10.1142/9789812833761_0009

Leitao, N., & Waugh, R. F. (2012). A Rasch measure of teachers' views of teacher-student relationships in the primary school. *Journal of Applied Measurement*, *13*(4), 403–427.

Lim, K. Y., & Lau, S. H. (2015). A structural model of mathematics achievement: Achievement, affective and demographic factors of students in lower secondary schools. *International Journal of Science and Mathematics Education*, *13*(5), 1075-1093. https://doi.org/10.1007/s10763-014-9564-4

Linacre, E. T. (1994). Estimating U.S. class a pan evaporation from few climate data. *Water International*, *19*(1), 5–14. https://doi.org/10.1080/02508069408686189

Linacre, J. M. (2003). *A user's guide to WINSTEPS: Rasch-model computer programs*. MESA Press.

Linacre, J. M. (2007). *Realiability and Separations: A user’s guide to Winsteps/Ministep Rasch– Model computer programs*. Winsteps.

Lins, R., & Kaput, J. (2004). The early development of algebraic reasoning: The current state of the field. In K. Stacey, H. Chick, & M. Kendal (Eds.), *The Future of the Teaching and Learning of Algebra: The 12th ICMI Study* (pp. 45-70). The University of Melbourne.

Mason, J. (1989). Mathematical abstraction as the result of a delicate shift of attention. *For the Learning of Mathematics*, *9*(2), 2–8. https://www.jstor.org/stable/40247947

Moss, J., London McNab, S. (2011). An Approach to Geometric and Numeric Patterning that Fosters Second Grade Students’ Reasoning and Generalizing about Functions and Co-variation. In Cai, J., Knuth, E. (eds), *Early Algebraization. Advances in Mathematics Education* (pp 277–301). Springer. https://doi.org/10.1007/978-3-642-17735-4_16

Mullis, I. V. S., & Martin, M. O. (Eds.). (2017). *TIMSS 2019 Assessment Frameworks*. Boston College, TIMSS & PIRLS International Study Center. http://timssandpirls.bc.edu/timss2019/frameworks/

Nathan, M. J., & Koedinger, K. R. (2000). An investigation of teachers’ beliefs of students’ algebra development. *Cognition and Instruction*, *18*(2), 209–237. https://doi.org/10.1207/s1532690xci1802_03

National Council of Teachers of Mathematics [NCTM]. (2000). *Principles and standards for school mathematics: An overview.* The National Council of Teachers of Mathematics.

Nickerson, R. S. (1985). Understanding understanding. *American Journal of Education*, *93*(2), 201–239. https://doi.org/10.1086/443791

Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what’s being reported? critique and recommendations. *Research in Nursing & Health*, *29*(5), 489–497. https://doi.org/10.1002/nur.20147

Polit, D. F., & Beck, C. (2017). *Nursing research: Generating and assessing evidence for nursing practice* (10th Eds.). Wolters Kluwer Health.

Ralston, N.C. (2013). *The development and validation of a diagnostic assessment of algebraic thinking skills for students in the elementary grades* (ED558693). [Doctoral dissertation, University of Washington]. ProQuest Dissertations and Theses Global.

Ralston, N. C., Li, M., & Taylor, C. (2018). The development and initial validation of an assessment of algebraic thinking for students in the elementary grades. *Educational Assessment*, *23*(3), 211–227. https://doi.org/10.1080/10627197.2018.1483191

Rittle-Johnson, B., & Alibali, M. W. (1999). Conceptual and procedural knowledge of mathematics: Does one lead to the other? *Journal of Educational Psychology*, *91*(1), 175–189. https://doi.org/10.1037/0022-0663.91.1.175

Saundarajan, K., Osman, S., Kumar, J. A., Daud, M. F., Abu, M. S., & Pairan, M. R. (2020). Learning algebra using augmented reality: A preliminary investigation on the application of photomath for lower secondary education. *International Journal of Emerging Technologies in Learning*, *15*(16), 123. https://doi.org/10.3991/ijet.v15i16.10540

Sfard, A., Linchevski, L. (1994). The gains and the pitfalls of reification - the case of algebra. In Cobb, P. (Eds), *Learning Mathematics* (pp.87–124). Springer. https://doi.org/10.1007/978-94-017-2057-1_4

Shea, T. L., Tennant, A., & Pallant, J. F. (2009). Rasch model analysis of the depression, anxiety and stress scales (DASS). *BMC Psychiatry*, *9, 2*1. https://doi.org/10.1186/1471-244x-9-21

Sibgatullin, I. R., Korzhuev, A. V., Khairullina, E. R., Sadykova, A. R., Baturina, R. V., & Chauzova, V. (2022). A systematic review on algebraic thinking in education. *Eurasia Journal of Mathematics, Science and Technology Education*, *18*(1), Article em2065. https://doi.org/10.29333/ejmste/11486

Smith, R. O. (2014). Beyond Passive Learning: Problem-based learning and concept maps to promote basic and higher-order thinking in basic skills instruction. *Journal of Research and Practice for Adult Literacy, Secondary, and Basic Education*, *3*(2), 50–55. https://eric.ed.gov/?id=EJ1045496

Somasundram, P. (2021). The role of cognitive factors in year five pupils' algebraic thinking: A structural equation modelling analysis. *Eurasia Journal of Mathematics, Science and Technology Education*, *17*(1), Article em1935. https://doi.org/10.29333/ejmste/9612

Stephens, A. C. (2008). What "counts" as algebra in the eyes of preservice elementary teachers? *The Journal of Mathematical Behavior*, *27*(1), 33–47. https://doi.org/10.1016/j.jmathb.2007.12.002

Sumintono, B., & Widhiarso, W. (2015). *Aplikasi pemodelan RASCH pada assessment pendidikan,* [Application of Rasch modeling in educational assessment]. Trim Komunikata.

Tofade, T., Elsner, J., & Haines, S. T. (2013). Best practice strategies for effective use of questions as a teaching tool. *American Journal of Pharmaceutical Education*, *77*(7), 155. https://doi.org/10.5688/ajpe777155

Usiskin, Z. (1988). Conceptions of school algebra and uses of variables. *The ideas of algebra, K-12*, 8-19. https://www.msri.org/attachments/workshops/454/Usiskin-conceptions%20of%20School%20Algebra.pdf

Van Amerom, B. A. (2002). *Reinvention of early algebra: Developmental research on the transition from arithmetic to algebra* (Doctoral dissertation, Freudenthal Institute).

Received: *February 24, 2023* Revised: *March 24, 2023* Accepted: *April 02, 2023*

Cite as: Jahudin, J., & Siew, N. M. (2023). An algebraic thinking skill test in problem-solving for seventh graders. *Problems of Education in the 21st Century*, *81*(2), 223-243. https://doi.org/10.33225/pec/23.81.223

Janet Jahudin	MEd. (Curriculum & Instructional Design), PhD Candidate, Faculty of Psychology and Education, University Malaysia Sabah, Jalan UMS, 88400, Kota Kinabalu, Sabah, Malaysia. E-mail: janetjahudin@gmail.com ORCID: https://orcid.org/0009-0000-9084-1150
Nyet Moi Siew *(Corresponding author)*	PhD, Senior Lecturer, Faculty of Psychology and Education, University Malaysia Sabah, Jalan UMS, 88400, Kota Kinabalu, Sabah, Malaysia. E-mail: sopiah@ums.edu.my ORCID: https://orcid.org/0000-0002-0937-9729