AKSIOMA: Jurnal Program Studi Pendidikan Matematika
Volume 12, No. 4, 2023, 3542-3560

ISSN 2089-8703 (Print)
ISSN 2442-5419 (Online)

DOI: https://doi.org/10.24127/ajpm.v12i4.7613

RASCH MODEL ANALYSIS: DEVELOPMENT OF HOTS-BASED MATHEMATICAL ABSTRACTION ABILITY INSTRUMENT ACCORDING TO RIAU ISLANDS CULTURE

Dwi Rismi Ocy[1], Wardani Rahayu[2], Makmuri[3]

[1,2,3] Universitas Negeri Jakarta, Jakarta, Indonesia

**Corresponding author. Universitas Negeri Jakarta, 13220, Jakarta, Indonesia*

E-mail: dwirismiocy@gmail.com [1)]
wardani.rahayu@unj.ac.id [2)]
makmuri@unj.ac.id [3*)]

Received 26 April 2023; Received in revised form 03 November 2023; Accepted 16 November 2023

Abstrak

Instrumen tes kemampuan abstraksi matematis berbasis HOTS sesuai budaya Kepulauan Riau merupakan bentuk soal yang masih jarang ditemukan dalam pembelajaran di sekolah. Tujuan penelitian ini adalah untuk menganalisis instrumen pengukur kemampuan abstraksi matematis berbasis HOTS sesuai budaya Kepulauan Riau untuk siswa SMA kelas 10. Responden sebanyak 402 siswa yang diambil dari empat sekolah menengah atas yang ada di kota Tanjungpinang. Data penelitian dikumpulkan dengan menggunakan tes pilihan ganda. Instrumen tes terdiri dari 19 butir soal yang diujicobakan. Tingkat kesulitan soal HOTS yang tinggi dan adanya keterbatasan waktu siswa dalam mengerjakan soal menyebabkan instrumen tes dibagi menjadi dua paket. Instrumen paket A sebanyak 10 butir, dan paket B sebanyak 9 butir. Analisis data menggunakan *software Winsteps*. Metode yang digunakan dalam analisis ini yaitu pemodelan Rasch. Hasil analisis instrumen paket A dan paket B dengan pemodelan Rasch diperoleh reliabilitas dengan nilai Alpha Cronbach berturut-turut sebesar 0,73 dan 0,75, yang berarti instrumen tes dapat diandalkan untuk mengukur kemampuan abstraksi matematis. Nilai *separation* paket A dan paket B sebesar 7,35 dan 6,15 menunjukkan item soal mempunyai sebaran respon *excellent*. Hasil analisis data menggunakan pemodelan Rasch menunjukkan seluruh item instrumen tes memenuhi kriteria *item fit order*. Berdasarkan hasil penelitian, 19 butir instrumen tes dikategorikan sangat layak untuk digunakan. Hasil penelitian pengembangan instrumen kemampuan abstraksi matematis berbasis HOTS sesuai budaya Kepulauan Riau memberikan kontribusi kepada siswa dalam pendekatan pendidikan matematika yang lebih inklusif dan menarik.

Kata kunci: Analisis model Rasch; HOTS; kemampuan abstraksi matematis; pengembangan instrumen.

Abstract

The HOTS-based mathematical abstraction ability test instrument according to the culture of the Riau Islands is a form of questions that is still rarely found in school learning. The goal of this study is to analyze the HOTS-based mathematical abstraction ability instrument according to the culture of the Riau Islands for 10th grade high school students. The respondents were 402 students taken from four high schools in the city of Tanjungpinang. Research data was collected using multiple choice questions. The test instrument consisted of 19 items that were tested. The high level of difficulty of the HOTS questions and the students' limited time in working on the questions caused the test instruments to be divided into two packages. Package A consists of 10 items, and package B consists of 9 items. Winsteps software is used for data analysis. Rasch modeling is a method applied in this analysis. Both package A and package B instruments were analyzed using Rasch modeling, the results showed reliability with Cronbach's Alpha values of 0.73 and 0.75 respectively, which means both instruments can be relied upon to measure students' mathematical abstraction abilities. The separation values of packages A and B are 7.35 and 6.15 respectively, which means the item have an excellent distribution of respons. Rasch modeling data analysis results demonstrate that every test instrument item satisfies the item fit order criteria. Based on research findings, 19 test instrument items were classified as extremely suitable for utilization. The development of a HOTS-based test instrument for mathematical abstraction ability that takes into account the cultural norms of the Riau Islands is helping students by promoting an even more engaging and inclusive mathematics education.

Keywords*: HOTS; instrument development; mathematical abstraction ability; Rasch model analysis.*

DOI: https://doi.org/10.24127/ajpm.v12i4.7613

INTRODUCTION

Students today must possess critical, analytical, logical, systematic, and creative thinking abilities as a result of the rapid advancement of life. These skills are needed to come up with answers for a wide range of present and future societal issues (Kristanto & Setiawan, 2020; Lestari, 2019; Nisa, Widyastuti, & Hamid, 2018). Critical, analytical, logical, systematic, and creative thinking abilities are related to the qualities of high-order thinking skills or HOTS. HOTS are necessary to meet upcoming difficulties. In order to compete with students from other countries, HOTS abilities are becoming more and more crucial for Indonesian high school students in the current era of globalization and the industrial revolution 4.0.

Mathematical abstraction ability plays a significant role in developing HOTS (Brookhart, 2010; Cetin & Dubinsky, 2017; Ocy, Rahayu, & Makmuri, 2023). When individuals engage in mathematical abstraction, they are required to identify patterns, make connections between different mathematical concepts, and apply these generalized principles to solve complex problems (KILIÇOĞLU & Kaplan, 2019; Murtianto, Sutrisno, Nizaruddin, & Muhtarom, 2019; Yusepa, 2017). This process fosters the development of higher order thinking skills such as analysis, synthesis, evaluation, and creativity. One of the key aspects of HOTS is the ability to analyze information critically and apply it in novel situations (Kurniasi & Arsisari, 2020; Setyaningsih & Mukodimah, 2022). Mathematical abstraction provides individuals with the capacity to deconstruct complex problems into simpler components, identify underlying patterns or principles, and then reconstruct this information to solve new problems (Annas, Djadir, & Hasma, 2018; Cetin & Dubinsky, 2017; Fitriani, Suryadi, & Darhim, 2018; Nisa', 2019; Simon, 2020; Sitihanifah & Ramlah, 2021; Syarifudin, Ratnaningsih, & Ni'mah, 2021). This process aligns with the skills required for higher order thinking, as it involves not only understanding mathematical concepts but also applying them in innovative ways.

According to research by Ocy et al. (2021), which examined students' HOTS abilities several schools in the Riau Islands Province. In other words, 79.5% of students were unable to abstract a problem so that the problem could be solved mathematically. This shows that the ability of high school students in Tanjungpinang is low in HOTS based mathematical abstraction ability.

If supported by appropriate resources, an effort to increase students' HOTS abilities in the area of mathematical abstraction will be successful, but 53.6% of teachers find it difficult to create engaging and contextual stimuli, and 51.8% of teachers find it difficult to create questions that adhere to HOTS criteria (Ocy, Rahayu, & Makmuri, 2021). The need for an instrument that can help teachers motivate and gauge the level of their students' HOTS talents in mathematical abstraction is where the urgency of this research lies. However, it has been discovered that the exploratory nature of the current study adds its significance, while the majority of implementations using the abstraction process have been qualitative (Hutagalung, Mulyana, & Pangaribuan, 2020; Komala, 2018; Murtianto et al., 2019; Nurrahmah, Zaenuri, & Wardono, 2021; Pratidiana,

ISSN 2089-8703 (Print)
ISSN 2442-5419 (Online)

DOI: https://doi.org/10.24127/ajpm.v12i4.7613

Rifa'i, & Priyani, 2021; Simon, 2020; Ulia, Waluya, & Walid, 2022) and there have been very few experimental studies undertaken. This means that this research will clearly contribute to the literature. In addition, a test instrument developed based on HOTS-based mathematical abstraction ability indicators was used in this research. This kind of study is valuable because it provides a wealth of data about the efficiency and credibility of the test instrument.

In the age of modernization, many students of the multicultural nation of Indonesia today exhibit indifference to traditional values. In order to address this, creating culturally-based HOTS instruments is a solution to developing youths' appreciation of their culture (Kamid, Saputri, & Hariyadi, 2021; Khoriyah & Oktiningrum, 2021; Yuliani, Alfarisa, & Tiurlina, 2022). It is intended that by incorporating aspects of Riau Islands culture into the instrument's creation, students will be exposed to new stimuli. This fits the criterion for issues related to HOTS because the problem is one that students have never encountered before (Alfiatin & Oktiningrum, 2019; Kamid et al., 2021; Khoriyah & Oktiningrum, 2021).

In contrast to previous instrument development research that employs the Classical Test Theory (CTT) (Arifin & Retnawati, 2015, 2017; Kurniasi & Arsisari, 2020; Lestari, 2019; Masitoh & Aedi, 2020), this study produces items with high item qualifications through data analysis utilizing the Rasch model. A Rasch modeling analysis with help from Winsteps software will be used in order to understand how well students are performing on tasks, as well as how difficult the tasks are, and to evaluate the HOTS-based mathematical abstraction ability instrument that has been constructed in accordance with the Riau Islands' culture. Identifying the viability of the developed instrument is the aim of this study.

METHODS

This study is an R&D (Research and Development) study. The 4D model was used as the development model for this study. The Define, Design, Develop, and Disseminate phases were used to create the 4D model, which was created in 1974 by S. Thiagarajan, Dorothy S. Semmel, and Melvyn I. Semmel (Sugiyono, 2010).

Examining research needs, doing literature reviews, and conducting early investigations are all done during the define stage. Making a test instrument development grid, Table of Specification, that corresponds with the research instrument development indicators is the task completed during the design stage. The task completed during the development stage is creating instrument items using a predefined grid, which is subsequently verified by experts, and conducting small-scale trials to ensure that the questions that have been developed can function properly. In the dissemination stage, the activities carried out are test instruments that have been validated are tested in larger scale in four schools in Tanjungpinang.

This study used written exams with multiple-choice questions as a method of data gathering. Four schools, SMAN 1 Tanjungpinang, SMAN 2 Tanjungpinang, SMAS Maitreyawira, and SMAS Santa Maria, with a combined enrollment of 402 students of 10th graders, participated in the field exams. Table 1 displays all available information.

ISSN 2089-8703 (Print)
ISSN 2442-5419 (Online)

DOI: https://doi.org/10.24127/ajpm.v12i4.7613

Table 1. Field test subjects

Name of School	Class	Number of students
SMAN 1 Tanjungpinang	X.3	44
	X.4	42
	X.5	41
	X.6	43
SMAN 2 Tanjungpinang	X.4	42
	X.7	41
	X.8	43
	X.9	42
SMAS Maitreyawira	X	33
SMAS Santa Maria	X	31
Total		**402**

The statistics fit test and instrument psychometric validity were both used as data analysis techniques in this study. As for the development of test instruments on trigonometry comparison material, there are 19 items that have been constructed and produced were empirically examined in this study. The 19 items were split into two question packages before the test, Package A and Package B, each of which contained ten and nine questions, respectively. This was done because the results of the small-scale test at the development stage were that the number of questions was large (19 items) and the HOTS level questions (C4, C5, C6) tended to have a high level of difficulty, considering that the time provided by the school was only two hours of lessons (100 minutes), students were unable to complete all the questions (19 items). Bearing this in mind, the test instruments were separated into two packages. Mathematical experts have approved the test instruments, which are divided in accordance with the specifications table from the design stage and satisfy all indicators of mathematical abstraction ability based on HOTS criteria. Table 2 provides specifics on how each item in each package was distributed.

Table 2. Distribution of items in each package

Packet	Item Number	Number of items on each HOTS level			Number of items
		Analyze	Evaluate	Create	
A	1, 2, 3, 4, 10, 11, 14, 15, 16, 29	4	5	1	10
B	8, 9, 12, 17, 23, 24, 25, 26, 30	4	4	1	9

An analysis of the class is done prior to giving out package A and package B questions to each class. It's important to remember that students in this subject have to have completed HOTS coursework on trigonometric comparison materials. By using the Slovin formula to calculate the number of students who would become respondents in each school, the decision was made to test question packages A and B for ten classes using a proportional sampling technique. This resulted in 198 students for testing package A and 204 students for testing package B. Table 3 displays the full set of data.

Table 3. Distribution of questions packets

Name of School	Class	Number of students	Packet
SMAN 1 Tanjungpinang	X.3	44	B
	X.4	42	A
	X.5	41	A
	X.6	43	B
SMAN 2 Tanjungpinang	X.4	42	B
	X.7	41	A
	X.8	43	B
	X.9	42	A
SMAS Maitreyawira	X	16	A
		17	B
SMAS Santa Maria	X	16	A
		15	B

DOI: https://doi.org/10.24127/ajpm.v12i4.7613

Item Fit Order

In Rasch model analysis, the fit of items is assessed to determine how well they contribute to the measurement of the underlying construct. When conducting Rasch model analysis, the fit of items is typically evaluated using fit statistics such as infit and outfit mean square (MNSQ) statistics (J M Linacre, 2019). The order in which items are fitted in Rasch model analysis is determined by examining the fit statistics for each item (Bambang Sumintono, 2018). The Rasch model uses the following criteria to determine an item's quality (Chan, Looi, & Sumintono, 2021; Maryati, Prasetyo, Wilujeng, & Sumintono, 2019; Perera, Sumintono, & Jiang, 2018; Bambang Sumintono, 2018).

1. Outfit MNSQ (Mean Square) value: $0{,}5 < outfit\ MNSQ < 1{,}5$
2. Outfit ZSTD (Z-Standard) value: $-2{,}0 < ZSTD < +2{,}0$
3. Point Measure Correlation value: $0{,}4 < Pt\ Measude\ Corr < 0{,}85$

The items of instruments that do not meet the three aforementioned conditions are deemed “misfit” and must be replaced; however, if the items of instruments that do meet at least two of the requirements are still deemed to be “fit” or in decent condition (Bambang Sumintono, 2018). Table 4 lists the requirements for item fit as determined by the instrument quality rating system (Boone, Staver, & Yale, 2014; John M. Linacre, 2010).

Table 4. Outfit MNSQ criteria

Item model fit mean-square range extremes	Criteria
$0{,}5 - 1{,}5$	Good
$1{,}5 - 2$	Marginal
$> 2{,}0$	Poor

Table 4 explains that MNSQ>2,0 indicates disruption to the measurement system, this suggests that the item is not functioning well within the measurement model and may need further investigation or revision; 1,5<MNSQ<2,0 indicates it has no significance for measurement, these values indicate some degree of misfit, they may still be acceptable depending on the specific research context or purpose of measurement; 0,5<MNSQ< 1,5 these values suggest that the item is functioning well within the measurement model and contributes meaningfully to the overall measurement; MNSQ<0,5 indicates it is not useful for measurement even though it doesn't cause damage to the measurement system.

Item Difficulty Level

According to Table 5, the Rasch modeling divides the item difficulty into four groups based on the Measure (logit) and the logit item's Standard Deviation (SD) values (Boone et al., 2014; Maryati et al., 2019; Perera et al., 2018; Bambang Sumintono, 2018).

Table 5. The criteria for the level of difficulty of the items

Measure (logit)	Interpretation of Item Difficulties
Measure logit $< -$SD	Very easy
$-$SD $\leq$ Measure logit $< 0{,}00$	Easy
$0{,}00 \leq$ Measure logit $\leq$ SD	Difficult
Measure logit $>$ SD	Very difficult

Item Discriminant Power

The classification of Pt Measure Corr values (Bambang Sumintono, 2018) can be shown in Table 6 and provides insight into the discriminating power of the items.

DOI: https://doi.org/10.24127/ajpm.v12i4.7613

Table 6. Criteria for discriminant power of items

Point Measure Correlation Value	Interpretation
> 0,40	Very good
0,30 – 0,39	Good
0,20 – 0,29	Fair
0,00 – 0,19	Cannot discriminate
< 0,00	Requires inspection of items

When the Pt Measure Corr value is 1.0, it indicates that all participants in tests with low abilities gave erroneous answers while all test-takers with high abilities gave accurate answers to all of the questions. A negative Pt Measure Corr number, on the other hand, shows items that were misleading because participants with low ability were able to correctly answer the questions while participants with high ability inadvertently provided incorrect answers. According to Bambang Sumintono (2018), questions with negative correlation values must be examined to determine whether the items need to be updated, should be removed from the test, or the answer keys are invalid.

Unidimensionality

Utilize the results of the unidimensionality test to evaluate the Rasch model's validity. The Rasch model requires that the variables are unidimensional. The dimensionality map, or more specifically, the raw variance data collected from the Winstep software, or by taking a look at the "raw variance explain measure" value, either show the unidimensional requirement. In the Rasch model, the unidimensional limit must be at least 40% and preferably more. Table 7 lists the Rasch model's unidimensionality requirements (Perera et al., 2018; Bambang Sumintono, 2018).

Table 7. Unidimensionality criteria

'Observed' Score in Raw Variance Explained by Measures	Criteria
> 60%	Excellent
40% – 60%	Good
20% – 40%	Fair
≥ 20%	Minimum
< 20%	Poor
< 15%	Unexpected variance

The dimensionality map also shows the independence needs. "Unexplained variance in 1st-5th" is an excellent indication of the value of independence. Table 8 displays the criterion for unexplained variance (Boone et al., 2014; Perera et al., 2018).

Table 8. Unexplained variance criteria

Unexplained variance in 1st-5th of PCA of residuals	Criteria
< 3%	Excellent
3 – 5%	Very Good
5 – 10%	Good
10 – 15%	Fair
> 15%	Poor

Monotonicity

Subsequently, investigate items monotonicity characteristics. The ability range of the response group known as group invariation (monotonization) is from low to high. When something is invariant, it either has a monotonous growth in item qualities or is curved. The observed average column, where the values must tend to increase gradually or consistently from small to large, reveals the nature of the monotonization in winsteps (J M Linacre, 2019). The instrument is monotonous meaning that there are no confusing instruments.

Reliability

Reliability is another factor that influences the quality of the item. With a Rasch reliability value of zero to one,

ISSN 2089-8703 (Print)
ISSN 2442-5419 (Online)

DOI: https://doi.org/10.24127/ajpm.v12i4.7613

often known as Cronbach's Alpha, person and item reliability describes the stability of respondents and items (Duncan, Bode, Lai, & Perera, 2003). The acceptable range for the instrument's coefficient Alpha Cronbach of reliability is 0.70 to 0.99 (Mohamad, Sulaiman, Sern, Mohd, & Salleh, 2015). The criteria in Table 9 serve as the foundation for establishing the reliability of both persons and items (J M Linacre, 2019; Perera et al., 2018).

Table 9. Person and item reliability

Person and Item Reliability	Criteria
$> 0{,}94$	Excellent
$0{,}91 - 0{,}94$	Very Good
$0{,}81 - 0{,}90$	Good
$0{,}67 - 0{,}80$	Fair
$< 0{,}67$	Poor

Person and Item Separation Index

The index of separation of persons and items estimates the tool that can distinguish between students' skills. The wider the spread of items from easy to difficult items, and the higher the index of separation of people and items, the more accurately the distribution of items responds to items (John M. Linacre, 2010). There are values for the separation index ranging from 0 to infinity, the higher the separation the better. Table 10 displays the criteria for the person index and item separation (Boone et al., 2014; Chan et al., 2021; J M Linacre, 2019).

Table 10. Person and item strata separation

Person and Item Strata Separation	Criteria
> 5	Excellent
$3 - 4$	Very Good
$2 - 3$	Good
$\geq 1{,}5$	Acceptable
$< 1{,}5$	Not Acceptable

Precision of Measurement

The precision of measurement is heavily dependent on the instrument, and it describes the result. To assess the adaptability and dependability of an instrument, accurate measurements are crucial. Less than 0.5 is the minimum acceptable standard of error for an instrument. In Rasch model, the estimated value of the items can be found in the "Model S.E" column (Perera et al., 2018).

Discriminating Power

The capacity of students who are able to answer questions and those who are unable to answer questions is separated using the discriminating power of questions. Analysis is employed in the Rasch modeling process as a method to assess discriminatory power at the level of the individual. To further identify groups of respondents, a method based on the respondent separation index can be employed. The higher the item separation value, the better the instrument is for all respondents and item items since it can differentiate between groups of items and groups of respondents (B Sumintono, 2016). The strata separation equation (H) is an additional equation to determine more specific grouping (Misbach & Sumintono, 2014).

$$H = \frac{[(4 \times SEPARATION) + 1]}{3} \qquad (1)$$

Students' HOTS in the Aspect of Mathematical Abstraction Ability

The examination of the Rasch model reveals students' capacities for mathematical abstraction. precisely by examining the logit score measure's mean value on the output of the person measure (Perera et al., 2018; B Sumintono, 2016). If the mean logit

ISSN 2089-8703 (Print)
ISSN 2442-5419 (Online)

DOI: https://doi.org/10.24127/ajpm.v12i4.7613

value is less than 0.00, then students have low HOTS abilities in the domain of mathematical abstraction. However, if the mean logit value is greater than 0.00, students are considered to have a strong HOTS ability in the field of mathematical abstraction.

RESULTS AND DISCUSSION

STAGE 1: DEFINE

In this study, three cognitive levels are included in the HOTS items: analyze (C4), evaluate (C5), and create (C6) (Phil, Kim, Zulnaidi, Syrene, & Rahim, 2022; Setyaningsih & Mukodimah, 2022; Tanujaya, 2016). Mathematical abstraction ability refers to the capacity to understand and manipulate abstract mathematical concepts, such as numbers, functions, and geometric shapes, without relying solely on concrete objects or real-world examples. While abilities include the level of mathematical abstraction, namely perceptual abstraction, internalization, interiorization and the second level of interiorization (Hong & Kim, 2016). The research indicators based on Hong & Kim (2016) research can be seen from Table 11.

Table 11. Mathematical abstraction ability indicator

Abstraction Level	Indicator
Perceptual Abstraction	Use physical things to learn about the qualities of mathematical objects. Find out about past experiences related to the current problem.
Internalization	utilizing mathematical terms, symbols, or pictures (such as graphs, diagrams, or other visual representations) to convey the findings of cognition. able to manipulate or solve difficulties.
Interiorization	assembling, compiling, developing, and organizing ideas to create new understandings or information.
Second Level of Interiorization	Use newly acquired knowledge broadly in a range of contexts.

STAGE 2: DESIGN

The first step in the design stage involves clearly defining the learning objectives that the tests aims to measure. This includes identifying the specific higher-order thinking skills that are targeted. In order to compile tests based on the hierarchical level of ability of HOTS, a table of specifications has been produced. At this point, the instrument design process starts, with three mathematics experts acting as the instrument's validators. The example of the table of specifications can be seen in Table 12.

Table 12. Example of the table of specifications for perceptual abstraction level

Abstraction Level	Indicator	HOTS Level	Item Number
Perceptual Abstraction	Students are able to recognize the properties of mathematical objects by utilizing physical objects related to trigonometric comparisons	C4 C5 C6	3, 23 - -
	Students are able to reacquaint themselves with previous experiences related to the problems they are facing regarding trigonometric comparisons	C4 C5 C6	4, 24 - -

DOI: https://doi.org/10.24127/ajpm.v12i4.7613

STAGE 3: DEVELOPMENT

At this stage, questions had been developed and constructed in accordance with the specifications table. There are currently twenty-five developed questions. Experts and panelists then vetted these 25 questions. Out of all the questions, only 19 were deemed valid by experts and panelists. The 19 questions that had passed the validation stage were then tested on a small scale. This is done to ensure that the questions do not contain writing errors, language that is difficult to understand, or calculation errors. Based on the results of the small-scale test, revision of the question items was carried out based on sentence formulation, completeness of stimulus information, inappropriate indicators, distribution of answer choices (distractors) that were less functional for multiple choice test questions.

STAGE 4: DISSEMINATION

Person Fit

Finding the appropriate participant for the Rasch model was the first stage in this study's data processing. According to the findings of the fit person analysis, which involved 402 respondents, 134 out of 198 respondents to the package A instrument and 128 out of 204 respondents to the package B instrument were found to be fit. Table 13 displays the findings of the analysis of misfit individuals.

Table 13. Misfit person

Misfit	Total	Person Number
Person packet A	64	005, 006, 009, 011, 014, 031, 034, 037, 038, 039, 040, 041, 044, 046, 047, 048, 071, 072, 073, 074, 075, 077, 080, 081, 082, 086, 088, 089, 090, 093, 094, 111, 112, 113, 115, 116, 118, 121, 122, 123, 128, 129, 130, 131, 132, 135, 152, 155, 157, 158, 159, 162, 163, 165, 166, 177, 178, 179, 190, 192, 194, 195, 196.
Person packet B	76	005, 006, 007, 009, 010, 011, 012, 013, 014, 015, 031, 041, 043, 048, 049, 050, 052, 053, 054, 055, 056, 068, 073, 076, 082, 083, 085, 090, 091, 094, 095, 097, 098, 099, 100, 101, 102, 103, 104, 106, 116, 121, 122, 123, 126, 127, 128, 133, 134, 135, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 152, 159, 163, 169, 173, 174, 180, 185, 186, 193, 194, 195, 200, 201, 202

In Rasch model analysis, Table 13, a misfit person refers to an individual whose responses to the items in a measurement instrument do not conform well to the expectations of the model. A misfit person in Rasch model analysis refers to someone whose responses deviate significantly from what would be expected based on their ability level. This deviation can manifest as consistently endorsing items that are too easy or too difficult for their ability level or showing inconsistent response patterns across items.

Identifying misfit persons is important because their responses can introduce noise and bias into the measurement process. Misfit persons may have different response patterns due to various reasons such as guessing, misunderstanding of items, or lack of motivation. Their inclusion in the analysis can compromise the validity and reliability of the measurement instrument.

Item Fit

After the outlier (misfit persons) is removed from the data, the data is once again examined. Being able to observe the item quality or item fit order information is valuable information from Rasch modeling. In Rasch model analysis, the fit of items is

AKSIOMA: Jurnal Program Studi Pendidikan Matematika
Volume 12, No. 4, 2023, 3542-3560
ISSN 2089-8703 (Print)
ISSN 2442-5419 (Online)

DOI: https://doi.org/10.24127/ajpm.v12i4.7613

assessed to determine how well they contribute to the measurement of the underlying construct. The fit statistics in Rasch analysis are used to evaluate whether the observed data conform to the expectations of the Rasch model. If the items function normally to the measurement or not, it is explained by item fit order. Items that are valid against the Rasch model are those that are normally functioning. The fit of items is crucial for ensuring that the measurement instrument is valid and reliable. Table 14 contains the findings from the fit order item analysis.

Table 14. Item fit order

Package	Entry Number	Item Number	Outfit			Point Measure Correlation	Point Measure Corr. Criteria	Fit/ Misfit
			MNSQ	MNSQ Criteria	ZSTD			
A	S9	16	0,52	Good	−0,63	0,42	Very good	Fit
	S10	29	0,51	Good	−0,66	0,43	Very good	Fit
	S1	1	0,75	Good	−0,73	0,56	Very good	Fit
	S6	11	0,70	Good	−0,91	0,57	Very good	Fit
	S5	10	0,69	Good	−0,99	0,59	Very good	Fit
	S2	2	0,75	Good	−0,80	0,57	Very good	Fit
	S4	4	0,88	Good	−0,21	0,55	Very good	Fit
	S7	14	0,96	Good	0,01	0,55	Very good	Fit
	S3	3	0,93	Good	−0,05	0,55	Very good	Fit
	S8	15	1,01	Good	0,17	0,54	Very good	Fit
B	S9	17	0,67	Good	−0,25	0,39	Good	Fit
	S1	8	0,75	Good	−0,54	0,51	Very good	Fit
	S2	9	0,73	Good	−0,73	0,56	Very good	Fit
	S8	30	0,65	Good	−1,14	0,60	Very good	Fit
	S7	26	0,79	Good	−0,63	0,57	Very good	Fit
	S5	12	0,69	Good	−1,45	0,71	Very good	Fit
	S4	24	1,19	Good	0,73	0,64	Very good	Fit
	S6	25	0,62	Good	−1,03	0,68	Very good	Fit
	S3	23	1,26	Good	0,66	0,48	Very good	Fit

Table 14 shows that every item, specifically the 19 items found in packages A and B of instruments, satisfies the very good fit item requirement. The Point Measure Correlation value with very good criteria shows that participants in tests are not confused when selecting answers to the instrument's questions, meaning that all participants with low abilities choose the wrong answers to the questions and all participants with high abilities choose the right answers.

Item Difficulty Level

To determine which items belong into the very easy to extremely difficult categories, the difficulty of each item is analyzed. The SD logit values for package A and B items are 1.95 and 1.65, respectively, according to the Rasch model's computations. It is possible to determine the difficulty level of each item based on the measure logit and SD logit values. The findings of the item difficulty level analysis can be summarized as shown in Table 15 based on the above explanation.

ISSN 2089-8703 (Print)
ISSN 2442-5419 (Online)

DOI: https://doi.org/10.24127/ajpm.v12i4.7613

Table 15. Item categories based on difficulty level

Instrument	Measure logit	Category	Item Number	Number of item
Packet A	$< -1,95$	Very easy	3, 4, 14, 15	4
	$-1,95 - 0$	Easy	-	0
	$0 - 1,95$	Difficult	1, 2, 10, 11	4
	$> 0,95$	Very difficult	16, 29	2
Packet B	$< -1,65$	Very easy	23,25	2
	$-1,65 - 0$	Easy	24	1
	$0 - 1,65$	Difficult	8, 9, 12, 26, 30	5
	$> 0,65$	Very difficult	17	1
Total items				**19**

Unidimensionality

Values from the unidimensionality test are used in the Rasch model analysis. Table 16 displays the raw variance information for packages A and B that was collected from the Winsteps software.

Table 16. Instrument unidimensionality

Raw variance explained by measures	Eigenvalue	Observed	Unidimensionality Criteria
Paket A	9,9866	50,0%	Good
Paket B	7,6244	45,9%	Good

The instruments in packages A and B are legitimate according to the unidimensional criteria of packages A and B, as shown by the value of "raw variance explained by measure" in the table. This demonstrates that packages A and B, which are HOTS-based instruments developed in accordance with the culture of the Riau Islands, can test students' ability in mathematical abstraction. Table 17 displays the independence values for packages A and B.

Table 17. Unexplained variance in 1^{st}-5^{th} of the instrument

Instrument	Unexplained variance	Eigenvalue	Observed	Criteria
Packet A	Unexplained variance in 1^{st} contrast	1,7489	8,8%	Good
	Unexplained variance in 2^{nd} contrast	1,4633	7,3%	Good
	Unexplained variance in 3^{rd} contrast	1,3103	6,6%	Good
	Unexplained variance in 4^{th} contrast	1,2140	6,1%	Good
	Unexplained variance in 5^{th} contrast	1,1874	5,9%	Good
Packet B	Unexplained variance in 1^{st} contrast	1,8422	11,1%	Fair
	Unexplained variance in 2^{nd} contrast	1,5336	9,2%	Good
	Unexplained variance in 3^{rd} contrast	1,3163	7,9%	Good
	Unexplained variance in 4^{th} contrast	1,1715	7,0%	Good
	Unexplained variance in 5^{th} contrast	1,0024	6,0%	Good

Based on Table 17, the instrument's items have a fairly high level of independence and satisfy the Rasch model's requirements.

Monotonicity

Positive sequential distance values are not separated in order to ascertain the type of monotonization, and it is stated that the answer categories can be

ISSN 2089-8703 (Print)
ISSN 2442-5419 (Online)

DOI: https://doi.org/10.24127/ajpm.v12i4.7613

viewed as an ordinal scale. The column observed average, where the values must tend to monotonically grow, illustrates the nature of the monotonicity in winsteps. View Figures 1 and 2.

```
---------------------------------------------------------------------------------------
|CATEGORY      OBSERVED|OBSVD SAMPLE|INFIT OUTFIT| COHERENCE |      |ESTIM|
|LABEL    SCORE COUNT %|AVRGE EXPECT|  MNSQ  MNSQ| M->C C->M | RMSR |DISCR|
|----------------------+------------+------------+-----------+------+-----|
|  0   0     741   55| -2.08 -2.08|  1.06   .94|  81%  87% | .3294|     | A
|  1   1     599   45|  1.59  1.58|  1.00   .63|  82%  74% | .3773| 1.01| B
---------------------------------------------------------------------------------------
```

Figure 1. Monotonicity of packet A

```
---------------------------------------------------------------------------------------
|CATEGORY      OBSERVED|OBSVD SAMPLE|INFIT OUTFIT| COHERENCE |      |ESTIM|
|LABEL    SCORE COUNT %|AVRGE EXPECT|  MNSQ  MNSQ| M->C C->M | RMSR |DISCR|
|----------------------+------------+------------+-----------+------+-----|
|  0   0     654   57| -1.86 -1.86|  1.06  1.02|  80%  82% | .3473|     | A
|  1   1     498   43|  1.39  1.39|   .99   .65|  77%  76% | .3877| 1.00| C
---------------------------------------------------------------------------------------
```

Figure 2. Monotonicity of packet B

The analysis reveals that in both package A and package B, the value in the observed average column has increased from a negative to a positive value. Both package A's and package B's observed average values tend to rise from -2.08 to 1.59 and from -1.86 to 1.39, respectively. The Root Mean Square Residual (RMSR) values for packages A and B grew as well, rising from 0.3294 to 0.3773 and 0.3473 to 0.3877, respectively. This demonstrates that the products utilized have complied with monotonicity condition.

Reliability

Rasch reliability can be used to examine the stability of persons and items of the instrument, which is another useful insight provided by Rasch modeling. Any reliability value around one is thought to be internally consistent. Figures 3 and 4 show the value of the person index and item reliability for packages A and B.

```
     SUMMARY OF 134 MEASURED (EXTREME AND NON-EXTREME) Person
---------------------------------------------------------------------------------
|          TOTAL                         MODEL         INFIT         OUTFIT    |
|          SCORE     COUNT     MEASURE    S.E.      MNSQ   ZSTD   MNSQ   ZSTD  |
|-------------------------------------------------------------------------------|
| MEAN       4.5      10.0        -.44     .94                                  |
|  SEM        .2        .0         .16     .01                                  |
| P.SD       2.3        .0        1.88     .15                                  |
| S.SD       2.3        .0        1.88     .15                                  |
| MAX.      10.0      10.0        4.81    1.91                                  |
| MIN.        .0      10.0       -4.80     .85                                  |
|-------------------------------------------------------------------------------|
| REAL RMSE   1.01 TRUE SD    1.58  SEPARATION  1.57  Person RELIABILITY  .71 |
|MODEL RMSE    .95 TRUE SD    1.62  SEPARATION  1.71  Person RELIABILITY  .74 |
| S.E. OF Person MEAN = .16                                                     |
---------------------------------------------------------------------------------
Person RAW SCORE-TO-MEASURE CORRELATION = 1.00 (approximate due to missing data
CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .73  SEM = 1.19 (a
STANDARDIZED (50 ITEM) RELIABILITY = .94

     SUMMARY OF 10 MEASURED (NON-EXTREME) Item
---------------------------------------------------------------------------------
|          TOTAL                         MODEL         INFIT         OUTFIT    |
|          SCORE     COUNT     MEASURE    S.E.      MNSQ   ZSTD   MNSQ   ZSTD  |
|-------------------------------------------------------------------------------|
| MEAN      59.9     134.0         .00     .26      1.02    .24    .77   -.48 |
|  SEM      11.5        .0         .65     .01       .02    .13    .06    .13 |
| P.SD      34.6        .0        1.95     .03       .05    .39    .17    .40 |
| S.SD      36.5        .0        2.05     .03       .05    .41    .17    .42 |
| MAX.     101.0     134.0        2.77     .32      1.09    .70   1.01    .17 |
| MIN.      14.0     134.0       -2.27     .23       .95   -.48    .51   -.99 |
|-------------------------------------------------------------------------------|
| REAL RMSE    .26 TRUE SD    1.93  SEPARATION  7.35  Item   RELIABILITY  .98 |
|MODEL RMSE    .26 TRUE SD    1.93  SEPARATION  7.48  Item   RELIABILITY  .98 |
| S.E. OF Item MEAN = .65                                                       |
---------------------------------------------------------------------------------
```

Figure 3. Person and item reliability packet A

ISSN 2089-8703 (Print)
ISSN 2442-5419 (Online)

DOI: https://doi.org/10.24127/ajpm.v12i4.7613

SUMMARY OF 128 MEASURED (EXTREME AND NON-EXTREME) Person

	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	3.9	9.0	-.56	1.00				
SEM	.2	.0	.18	.02				
P.SD	2.3	.0	2.00	.26				
S.SD	2.3	.0	2.00	.26				
MAX.	9.0	9.0	4.30	1.92				
MIN.	.0	9.0	-4.53	.82				

REAL RMSE 1.09 TRUE SD 1.68 SEPARATION 1.54 Person RELIABILITY .70
MODEL RMSE 1.03 TRUE SD 1.71 SEPARATION 1.66 Person RELIABILITY .73
S.E. OF Person MEAN = .18

Person RAW SCORE-TO-MEASURE CORRELATION = .99
CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .75 SEM = 1.16
STANDARDIZED (50 ITEM) RELIABILITY = .94

SUMMARY OF 9 MEASURED (NON-EXTREME) Item

	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	55.3	128.0	.00	.25	1.03	.18	.82	-.49
SEM	9.6	.0	.58	.01	.06	.45	.08	.25
P.SD	27.2	.0	1.65	.02	.18	1.28	.22	.71
S.SD	28.9	.0	1.75	.02	.19	1.35	.24	.76
MAX.	98.0	128.0	2.47	.30	1.50	3.24	1.26	.73
MIN.	16.0	128.0	-2.66	.23	.80	-1.78	.62	-1.45

REAL RMSE .27 TRUE SD 1.63 SEPARATION 6.16 Item RELIABILITY .97
MODEL RMSE .25 TRUE SD 1.63 SEPARATION 6.42 Item RELIABILITY .98
S.E. OF Item MEAN = .58

Figure 4. Person dan item reliability packet B

The red mark in Figure 3 shows that the Cronbach Alpha coefficient is 0.73, the item reliability is 0.98, and the person reliability index for package A is 0.71. The strong reliability estimate for package A shows how consistently items and student responses interact. Similar results can be seen in Figure 4, where the Cronbach Alpha coefficient is 0.75, item reliability is 0.97, and the person reliability index for package B is 0.70. As a result, package A and package B instruments are regarded to be reliable and having excellent psychometric internal consistency.

Person and Item Separation Index

The person separation index value for package A is 1.57 and the item separation index value is 7.35 in Figure 3. The person separation index value for package B is 1.54 and the item separation index value is 6.16 in Figure 4. This shows that the distribution range of respondents in packages A and B has person separation with "acceptable" criteria and item separation in packages A and B instruments with "very good" specifications. The person and item separation index indicates that the instrument respondents who meet the criteria can be relied upon and that the level of student ability has an excellent level of diversity.

Precision of Measurement

Precision of measurement is a strong reliance on the instrument and describes the conclusion. Figures 3 and 4 show that the S.E. model has a value of 0.26 and 0.25, respectively, which is less than 0.5. As a result, measurement accuracy serves as a trustworthy indicator of how well an item fits. The instrument's level of reliability was determined to be high, and it demonstrated accurate measurement.

Discriminatory Power

The values of item separation packages A and B are 7.35 and 6.16, respectively, according to Figures 3 and 4. Packages A and B had person separation values of 1.57 and 1.54, respectively. According to (Perera et al., 2018; Bambang Sumintono, 2018) stratum separation equation formula (H), specifically:

$$H_{item\ packet\ A} = \frac{[(4 \times 7{,}35) + 1]}{3} = 10{,}133 \cong 10$$

$$H_{person\ packet\ A} = \frac{[(4 \times 1{,}57) + 1]}{3} = 2{,}42 \cong 2$$

DOI: https://doi.org/10.24127/ajpm.v12i4.7613

$$H_{item\ packet\ B} = \frac{[(4 \times 6{,}16) + 1]}{3} = 8{,}546 \cong 9$$

$$H_{person\ packet\ B} = \frac{[(4 \times 1{,}54) + 1]}{3} = 2{,}386 \cong 2$$

The stratum equation's (H) value for the items in packages A and B turns out to be 10.133, which is rounded to 10, and 8.546, which is rounded to 9. As for the respondents in package A and B, the results obtained are 2.42 which is rounded up to 2 and 2.386, rounded up to 2, respectively. As a result, it is possible to identify 10 groups of items A and 9 groups of items B. For respondents, this means that package A and package B respondents can be sorted into two groups that represent the skills of the students.

Based on the results of the psychometric validity analysis and the fit statistics test, the results of the field test data analysis can be summarized in Table 18.

Table 18. Summary of field test data analysis

Number	Criteria	Packet A	Packet B
1	Unidimensionality	50% (Good)	45,9% (Good)
2	Monotonicity	−2,08 to 1.59 (Monoton)	−1,86 to 1,39 (Monoton)
3	Alpha Cronbach	0,73 (Adequate)	0,75 (Adequate)
4	Person Reliability	0,71 (Fair)	0,70 (Fair)
5	Item Reliability	0,98 (Excellent)	0,97 (Excellent)

The following findings can be obtained from Table 18, the examination of the instrument's psychometric validity: (1) measurement packages A and B were able to explain 50% and 45.9% of the raw variance, respectively, showing that they met the conditions for unidimensionality; (2) with scores for package A ranging from -2.08 to 1.59 and for package B from -1.86 to 1.39, the monotonization of packages A and B tended to rise (meeting the monotonicity criteria); (3) the A and B instruments' respective Alpha Cronbach reliability scores of 0.73 and 0.75 show that respondents' consistency in responding to the instrument can be accepted holistically; (4) the person reliability scores for packages A and B were 0.70 and 0.71, respectively, indicating that the subject's consistency in giving an answer proved quite good; (5) the item reliability scores for packages A and B were 0.97 and 0.98, respectively, indicating the excellence of package A and B instruments in the special category. All items examined empirically met valid and reliable criteria according to the findings of statistical fit test analysis and psychometric validity.

Students' HOTS in the Aspect of Mathematical Abstraction Ability

The examination of the Rasch model reveals students' capacities for mathematical abstraction. Precisely by examining the logit score measure's mean value on the output of the person measure. For respondents in package A, the mean value for the person measure is $-0.44 < 0.00$ logit. Additionally, package B respondents' mean value for the person measure is $-0{,}56 < 0{,}00$ logit. This demonstrates that students' mathematical abstraction skills fall below of the required level for the item's difficulty, which means that HOTS performance in the area of students' mathematical abstraction abilities is poor.

ISSN 2089-8703 (Print)
ISSN 2442-5419 (Online)

DOI: https://doi.org/10.24127/ajpm.v12i4.7613

DISCUSSION

The research findings revealed that Tanjungpinang high school students have low HOTS abilities in the area of mathematical abstraction. This demonstrates the necessity of improving students' HOTS skills, particularly in the area of mathematical abstraction. There are several strategies that a mathematics teacher can use to improve their students' mathematical abstraction abilities. The following is how this research contributes to students' stronger HOTS abilities in the area of mathematical abstraction ability: (1) Students can use the final product as training material to develop higher-order thinking skills in the mathematical abstraction ability area; (2) High school math teachers can use the final product to assess their students' understanding of mathematical concepts, aptitude for mathematical abstraction, and capacity for higher-order thinking.

To give students fresh insights, a cultural component—specifically, the culture of the Riau Islands—was incorporated into the instrument's construction. The problem environment that this instrument develops presents new challenges that students have never encountered before. That is one of the study's advantages. A further advantage of this research is that, in contrast to CTT, which focuses on observed scores and assumes equal item discrimination across different ability levels, the Rasch model operates on a more fundamental level by estimating item difficulty and person ability parameters independently. This distinction gives rise to several advantages that position the Rasch analysis model as a powerful tool for measurement and assessment.

CONCLUSION AND SUGGESTION

Based on the results of the research and discussion, the following conclusions were obtained: (1) The final product in this research produced a HOTS instrument on the aspect of mathematical abstraction abilities according to the culture of the Riau Islands. The test instrument is a set of multiple-choice questions consisting of 19 questions. The validity of the instrument is proven by the results of expert assessment and Rasch model analysis. The instrument also meets the criteria for reliability; (2) Multiple-choice questions have the characteristics of item difficulty levels that tend to be difficult or very difficult, good unidimensionality, fulfill monotonous characteristics, and very good precision of measurement.

This research's limitation is the instrument's incorporation of cultural components from the Riau Islands, which makes it challenging for students from other regions to use. Because students from other regions are unfamiliar with the cultural background of the questions. This can be overcome by conducting research regarding the development of HOTS instruments on aspects of mathematical abstraction abilities by involving other cultural contexts that are familiar and close to students.

REFERENCES

Alfiatin, A. L., & Oktiningrum, W. (2019). Pengembangan Soal Higher Order Thinking Skills Berbasis Budaya Jawa Timur Untuk Mengukur Penalaran Siswa Sd. *Indiktika : Jurnal Inovasi Pendidikan Matematika*, *2*(1), 30. https://doi.org/10.31851/indiktika.v2i1.3395

DOI: https://doi.org/10.24127/ajpm.v12i4.7613

Annas, S., Djadir, & Hasma, S. M. (2018). The Abstraction Ability in Constructing Relation Within Triangles by the Seventh Grade Students of Junior High School. *Journal of Physics: Conference Series*. https://doi.org/10.1088/1742-6596/954/1/012029

Arifin, Z., & Retnawati, H. (2015). Analisis Instrumen Pengukur Higher Order Thinking Skills (HOTS) Matematika Siswa SMA. *Seminar Nasional Matematika Dan Pendidikan Matematika UNY*, (20), 783–790.

Arifin, Z., & Retnawati, H. (2017). Pengembangan instrumen pengukur higher order thinking skills matematika siswa SMA kelas X. *PYTHAGORAS: Jurnal Pendidikan Matematika*. https://doi.org/10.21831/pg.v12i1.14058

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

Brookhart, S. M. (2010). How To Assess Higher Order Thinking Skills in Your Classroom. In *ASCD*.

Cetin, I., & Dubinsky, E. (2017). Reflective abstraction in computational thinking. *Journal of Mathematical Behavior*, *47*, 70–80. https://doi.org/10.1016/j.jmathb.2017.06.004

Chan, S. W., Looi, C. K., & Sumintono, B. (2021). Assessing computational thinking abilities among Singapore secondary students: a Rasch model measurement analysis. *Journal of Computers in Education*, *8*(2), 213–236. https://doi.org/10.1007/s40692-020-00177-2

Duncan, P. W., Bode, R. K., Lai, S. M., & Perera, S. (2003). Rasch Analysis of a New Stroke-Specific Outcome Scale. In *The Stroke Impact Scale* (pp. 950–963).

Fitriani, N., Suryadi, D., & Darhim, D. (2018). the Students' Mathematical Abstraction Ability Through Realistic Mathematics Education With Vba-Microsoft Excel. *Infinity Journal*, *7*(2), 123. https://doi.org/10.22460/infinity.v7i2.p123-132

Hong, J. Y., & Kim, M. K. (2016). Mathematical Abstraction in the Solving of Ill-Structured Problems by Elementary School Students in Korea. *EURASIA Journal of Mathematics, Science and Technology Education*, *12*(2), 267–281. https://doi.org/10.12973/eurasia.2016.1204a

Hutagalung, E. E., Mulyana, E., & Pangaribuan, T. R. (2020). Mathematical abstraction: Students' concept of triangles. *Journal of Physics: Conference Series*, *1521*(3). https://doi.org/10.1088/1742-6596/1521/3/032106

Kamid, K., Saputri, R., & Hariyadi, B. (2021). Pengembangan Soal Higher Order Thinking Skills Berbasis Budaya Jambi. *Jurnal Cendekia : Jurnal Pendidikan Matematika*, *5*(2), 1793–1806. https://doi.org/10.31004/cendekia.v5i2.678

Khoriyah, M., & Oktiningrum, W. (2021). Pengembangan Soal Higher Order Thinking Skills (HOTS) Berbasis Budaya Lokal Blitar untuk Mengukur Dimensi

DOI: https://doi.org/10.24127/ajpm.v12i4.7613

Pengetahuan Matematika Siswa kelas V Sekolah Dasar. *Bina Gogik: Jurnal Ilmiah Pendidikan Guru Sekolah Dasar*, *8*(1), 93–100. Retrieved from http://repository.uniramalang.ac.id/id/eprint/270%0A

KILIÇOĞLU, E., & Kaplan, A. (2019). An Examination of Middle School 7th Grade Students' Mathematical Abstraction Processes. *Journal of Computer and Education Research*, *7*(13), 233–256. https://doi.org/10.18009/jcer.547975

Komala, E. (2018). Analysis of Students' Mathematical Abstraction Ability By Using Discursive Approach Integrated Peer Instruction of Structure Algebra Ii. *Infinity Journal*, *7*(1), 25. https://doi.org/10.22460/infinity.v7i1.p25-34

Kristanto, P. D., & Setiawan, P. G. F. (2020). Pengembangan Soal HOTS (Higher Order Thinking Skills) Terkait Dengan Konteks Pedesaan. *PRISMA: Prosiding Seminar Nasional Matematika*, *3*, 370–376. Retrieved from https://journal.unnes.ac.id/sju/index.php/prisma/article/view/37616

Kurniasi, E. R., & Arsisari, A. (2020). Pengembangan Instrumen Pengukur Higher Order Thinking Skills (HOTS) Matematika Pada Siswa SMP. *AKSIOMA: Jurnal Program Studi Pendidikan Matematika*, *9*(4), 1213–1222. https://doi.org/http://dx.doi.org/10.24127/ajpm.v9i4.3162

Lestari, S. A. P. (2019). Pengembangan Instrumen Asesmen Higher Order Thinking Skill (HOTS) pada Materi Himpunan Kelas VII SMP. *JKPM (Jurnal Kajian Pendidikan Matematika)*, *4*(2), 111. https://doi.org/10.30998/jkpm.v4i2.3862

Linacre, J M. (2019). A User's Guide to WINSTEPS® Rasch-Model Computer Programs: Program Manual 4.4. 6. *Mesa-Press, Chicago, IL*.

Linacre, John M. (2010). Predicting responses from rasch measures. *Journal of Applied Measurement*, *11*(1), 1–10.

Maryati, Prasetyo, Z. K., Wilujeng, I., & Sumintono, B. (2019). Measuring teachers' pedagogical content knowledge using many-facet rasch model. *Cakrawala Pendidikan*, *38*(3), 452–464. https://doi.org/10.21831/cp.v38i3.26598

Masitoh, L. F., & Aedi, W. G. (2020). Pengembangan Instrumen Asesmen Higher Order Thinking Skills (HOTS) Matematika di SMP Kelas VII. *Jurnal Cendekia : Jurnal Pendidikan Matematika*, *4*(2), 886–897. https://doi.org/10.31004/cendekia.v4i2.328

Misbach, I. H., & Sumintono, B. (2014). Pengembangan dan Validasi Instrumen "Persepsi Siswa Tehadap Karakter Moral Guru" di Indonesia dengan Model Rasch. *PROCEEDING Seminar Nasional Psikometri*, 148–162.

Mohamad, M. M., Sulaiman, N. L., Sern, L. C., Mohd, K., & Salleh. (2015). *Measuring the Validity and Reliability of Research Instruments*.

Murtianto, Y. H., Sutrisno, Nizaruddin, & Muhtarom. (2019). Effect of Learning Using Mathematica Software Toward Mathematical Abstraction Ability, Motivation, and Independence of Students in

DOI: https://doi.org/10.24127/ajpm.v12i4.7613

Analytic Geometry. *Infinity Journal*, *8*(2), 219–228. https://doi.org/10.22460/infinity.v8i2.p219-228

Nisa', A. L. (2019). Analisis Kemampuan Abstraksi Matematis Siswa Dalam Menyelesaikan Soal Pada Materi Segiempat Kelas Vii Smp. *JPM : Jurnal Pendidikan Matematika*, *4*(1), 01. https://doi.org/10.33474/jpm.v4i1.2610

Nisa, N. A. K., Widyastuti, R., & Hamid, A. (2018). Pengembangan Instrumen Assesment Higher Order Thinking Skill (HOTS) Pada Lembar Kerja Peserta Didik Kelas VII SMP. *Prosiding Seminar Nasional Matematika Dan Pendidikan Matematika UIN Raden Intan Lampung*, *1*(2), 543–556.

Nurrahmah, A., Zaenuri, & Wardono. (2021). Analysis of students difficulties in mathematical abstraction thinking in the mathematics statistic course. *Journal of Physics: Conference Series*, *1918*(4). https://doi.org/10.1088/1742-6596/1918/4/042112

Ocy, D. R., Rahayu, W., & Makmuri, M. (2021). Experiences of Senior High School Mathematics Teachers With Higher Order Thinking Skills (HOTS) In Riau Islands Province-Indonesia. *Jurnal Gantang*, *VI*(2), 103–111.

Ocy, D. R., Rahayu, W., & Makmuri, M. (2023). Development of a HOTS-Based Mathematical Abstraction Ability Instrument in Trigonometry Using Riau Islands Province Culture. *Jurnal Gantang*, *8*(1), 37–52. https://doi.org/10.31629/jg.v8i1.5654

Perera, C. J., Sumintono, B., & Jiang, N. (2018). The Psychometric Validation of The Principal Practices Questionnaire Based on Item Response Theory. *International Online Journal of Educational Leadership*, *2*(1), 21–38.

Phil, R., Kim, T., Zulnaidi, H., Syrene, S., & Rahim, A. (2022). Development of Higher-Order Thinking Skills test instrument on Quadratic Equations (HOTS-QE) for Secondary School Students. *Pegem Journal of Education and Instruction*, *13*(1), 379–394. https://doi.org/10.47750/pegegog.13.01.41

Pratidiana, D., Rifa'i, R., & Priyani, D. (2021). Analysis of Mathematical Abstraction Ability of Grade Viii Students in Cartesian Coordinate Material. *MaPan*, *9*(1), 178. https://doi.org/10.24252/mapan.2021v9n1a12

Setyaningsih, N., & Mukodimah, T. D. (2022). Pengembangan Instrumen Asesmen Higher Order Thinking Skills (Hots) Berbasis Literasi Matematika Pada Materi SPLDV. *AKSIOMA: Jurnal Program Studi Pendidikan Matematika*, *11*(3), 1739–1748.

Simon, M. A. (2020). Elaborating reflective abstraction for instructional design in mathematics: Postulating a Second Type of Reflective Abstraction. *Mathematical Thinking and Learning*, *22*(2), 162–171. https://doi.org/10.1080/10986065.2020.1706217

Sitihanifah, N., & Ramlah. (2021). Analisis Kemampuan Abstraksi Matematis Siswa Kelas VIII pada Materi Segitiga. *MAJU : Jurnal*

ISSN 2089-8703 (Print)
ISSN 2442-5419 (Online)

DOI: https://doi.org/10.24127/ajpm.v12i4.7613

Ilmiah Pendidikan Matematika, *8*(2), 168–174.

Sugiyono, P. D. (2010). *Metode Penelitian Kuantitatif Kualitatif dan R&D*. Bandung: CV. ALVABETA.

Sumintono, B. (2016). Penilaian Keterampilan Berpikir Tingkat Tinggi : Aplikasi Pemodelan Rasch pada Asesmen Pendidikan. *Seminar Nasional Pendidikan IPA, FKIP Jurusan PMIPA, Universitas Lambung Mangkurat*.

Sumintono, Bambang. (2018). Rasch Model Measurements as Tools in Assesment for Learning. *Proceedings of the 1st International Conference on Education Innovation (ICEI 2017)*, *173*(Icei 2017), 38–42. https://doi.org/10.2991/icei-17.2018.11

Syarifudin, M. T., Ratnaningsih, N., & Ni'mah, K. (2021). Analisis Kemampuan Abstraksi Matematis dalam Pembelajaran Matematika di MAN 1 Tasikmalaya. *MUST: Journal of Mathematics Education, Science and Technology*, *6*(2), 231. https://doi.org/10.30651/must.v6i2.7461

Tanujaya, B. (2016). Development of an Instrument to Measure Higher Order Thinking Skills in Senior High School Mathematics Instruction. *Journal of Education and Practice*. https://doi.org/10.21831/pg.v4i2.557

Ulia, N., Waluya, S. B., & Walid. (2022). The level of mathematics reflective abstraction of pre-service elementary school teachers. *AIP Conference Proceedings*, *2577*(1). https://doi.org/10.1063/5.0096152

Yuliani, A., Alfarisa, F., & Tiurlina, T. (2022). The development of mathematical HOTS questions based on banten culture. *International Journal of Trends in Mathematics Education Research*, *5*(1), 44–52. https://doi.org/10.33122/ijtmer.v5i1.112

Yusepa, B. (2017). Kemampuan Abstraksi Matematis Siswa Sekolah Menengah Pertama (Smp) Kls Viii. *Symmetry: Pasundan Journal of Research in Mathematics Learning and Education*, *I*. https://doi.org/10.23969/symmetry.v1i1.233