JTAM (Jurnal Teori dan Aplikasi Matematika)
http://journal.ummat.ac.id/index.php/jtam
p-ISSN 2597-7512 | e-ISSN 2614-1175
Vol. 6, No. 3, July 2022, pp. 511-529

Measuring Changes of Students' Statistical Reasoning Taught by Ethnomathematics Approach Assisted TinkerPlots: A Stacking Analysis Study

Rahmi Ramadhani[1,2], **Sahat Saragih**[2], **Rizki Maulida**[3], **Sinta Makito Br Simanjuntak**[4]
[1,4]Department of Informatics, Universitas Potensi Utama, Medan, Indonesia
[2]Postgraduate Program, Universitas Negeri Medan, Medan, Indonesia
[3]Department of Industrial Engineering, Universitas Potensi Utama, Medan, Indonesia
rahmiramadhani3@gmail.com[1], saragihpps@gmail.com[2], rizkimaulida24@gmail.com[3],
sintamakita@gmail.com[4]

Article History:
Received : 11-04-2022
Revised : 16-05-2022
Accepted: 20-05-2022
Online : 01-07-2022

Keywords:
Stacking Analysis;
Rasch Model;
Statistical Reasoning;
Ethnomathematics;
TinkerPlots.

ABSTRACT

This study aims to use the Rasch model stacking analysis technique to assess students' statistical reasoning abilities in descriptive statistics learning utilizing a TinkerPlots-assisted ethnomathematics approach in Nias cultural environment. This research is a quasi-experimental study that uses pre-and post-test control group designs. The stacking technique is used to examine how students' statistical reasoning abilities change in the presence of the intervention. The sample for this study is students in the 12th year at Gunungsitoli High School in the Nias Islands, Sumatera Utara region. Students are administered an exam that consists of five essay questions. Their responses are evaluated using a rubric that incorporates diagnostic criteria and a response certainty index. The data is analyzed using the Rasch Partial Credit Model with WINTSTEPS 4.5.5. The results indicate that students in the experimental group improved their statistical reasoning abilities more than students in the control group when they used TinkerPlots-assisted ethnomathematics in Nias cultural environment (the group with an ordinary learning approach). Along with the intervention, it is discovered that changes in student learning capacities are achievable in some situations due to pupils receiving the correct response or cheating. In other circumstances, pupils who undergo negative changes may respond inappropriately due to carelessness, boredom, or misconceptions. Additionally, these findings demonstrate that some children provide correct responses following intervention on certain items. According to the findings of this study, stacking analysis approaches are critical for describing each change in student abilities as a result of each person and each item.

https://doi.org/10.31764/jtam.v6i3.8375

◆

A. INTRODUCTION

Statistical reasoning ability is part of mathematical reasoning ability. Reasoning itself includes abilities such as observing or assessing a chain of arguments, carrying out a verifying process related to obtaining, and finding out the differences between the results of evidence carried out and other reasoning processes, as well as revealing basic ideas related to arguments or certain statements, and designing arguments or statements from informal to formal form (Gürbüz & Erdem, 2016; Mumu & Tanujaya, 2019). Mathematical reasoning is a process of thinking to make decisions through critical and logical thinking. Mathematical reasoning is related to mathematical proof and the logical process accompanying it (Kilpatrick et al., 2009).

Mathematical reasoning is the main focus for teachers to help students understand mathematical concepts and apply them in solving mathematical problems. Statistical reasoning ability is an important part in developing students' mathematical abilities since statistics have a significant role in understanding data which will later be applied in the development of Internet of Things (IoT) and Artificial Intelligence (AI)-based technologies (Yaqoob et al., 2019). The data understanding eventually leads to the development of analytical skills on data closely related to statistics. Statistical data analysis aims to filter massive information so that later it can conclude valid and significant information used as further recommendations regarding the development of a system (Setiawan & Sukoco, 2021).

Statistical reasoning ability is essential for high school students because it has an impact on the understanding process in other scientific fields (Sharma, 2017). Statistical reasoning ability is closely related to students who have good statistical understanding and reasoning skills which will provide convenience in carrying out statistical modelling (Aridor & Ben-Zvi, 2018b; Justice et al., 2018). In addition, statistical reasoning ability is also closely related to mathematical communication skills. It will be noticed when a problem presenting a description of a graph or table of statistical data is assigned to the students (Hughes et al., 2020). The research conducted by Lepak, et al (2018) finds that students with low mathematical communication skills would have difficulty in expressing mathematical problems based on mathematical literacy and doing mathematical reasoning related to the given problem. Im & Jitendra (2020) also studied 338 students who struggle with mathematical learning and discovered that students work to solve arithmetic problems that need mathematical reasoning ability. Additionally, Ramadhani and Evans (2022) claim that statistical reasoning ability is a mathematical talent that enables students to evaluate data, connect concepts, and explain statistical techniques. Additionally, statistical reasoning ability enables students to strengthen their grasp of the fundamental principles and logic and their practical capacity to select, create rapidly, and evaluate descriptive and inferential procedures (delMas, 2002; Garfield, 2002). Students who develop statistical reasoning abilities will develop a critical, analytical, responsible, and capable approach toward decision-making.

Some of the previous research findings show that statistical reasoning ability plays a crucial role in the development of other students' mathematical abilities. In line with these findings, another fact is obtained that students' statistical reasoning abilities are low. On the world scale, Indonesian students are only in the bottom 6 for the 2018 PISA (Programme for International Student Assessment) assessment, and Indonesian students did not achieve an average score in the 2015 TIMSS (The Trends in International Mathematics and Science Study) assessment (Indonesia did not participate in the 2019 TIMSS assessment). If we look at the national assessment, it is found that 75.7% of Indonesian students do not have good mathematical literacy skills, which can be seen from the basic level achievement obtained by students - Level 2. This is also in line with the fact that students' statistical reasoning abilities are low which is proven by the various distress of students in solving problems in statistical data. Mahdayani (2016) finds that only 8.3% of students are able to do the processing data and draw conclusions from the data presented. Furthermore, Mahdayani's findings (2016) also show that less than 50% of students are able to understand the statistical data provided and only 16.5% of students are able to do statistical data transformation. This finding is also supported by the achievement of student learning outcomes actualized from the assessment of the 2019 National Examination results which show that there is a decrease in the average score annually (Pusat Penilaian Pendidikan Kementerian Pendidikan dan Kebudayaan, 2022b). The findings are also supported by the low response process of students who answered correctly on statistical material compared to other mathematics materials in the 2019 National Examination (Pusat Penilaian Pendidikan Kementerian Pendidikan dan Kebudayaan, 2022a).

Based on the facts and findings obtained, it is necessary to conduct further analysis related to the causes of the students' low statistical reasoning ability and why statistics subject is difficult to understand for students. The impact of the students' low statistical reasoning ability is not only seen in the results of the national exam and the mapping of teaching materials that are difficult for students to understand but also relates to the ability of students to manage their individual learning abilities. Moreover, it leads to anxiety and confusion even though they only solve a simple numerical problem (Ching et al., 2020). The tendency of students to memorize equations is one of the causes of students experiencing difficulties in the statistical reasoning process (Nuralam & Gadeng, 2018). Another reason is that students are not familiar to solve problems that are close to their traditions and culture. Students have been given informal problems related to statistical problems but they are not close to their traditions and culture. This condition makes it difficult for students to understand and reason the informal problems that are asked and the analytical procedures in accordance with the questions presented.

Based on these, the integration of culture in learning mathematics, especially those related to learning descriptive statistics, needs to be carried out. Cultural integration in mathematics or often called ethnomathematics can be done using either cultural artifacts or cultural mentifacts. The informal problem presentations provided in the context of artifacts will help students to understand problems that are familiar in daily life. Mathematics actually cannot construct itself and is very dependent on aspects inherent in the culture in which humans develop and manage their lives (Risdiyanti & Prahmana, 2020; Utami et al., 2021). The integration of cultural aspects in learning mathematics helps students in constructing the knowledge gained through experience and utilizing it to construct new knowledge gained through the learning process. The concept of ethnomathematics is related to the motive where a certain culture (Ethno) actualizes the steps to calculate, conclude, compare, and classify a technique and idea that enables people in that culture to model the environment context nature, the social context in mathematical phenomena (D'Ambrosio, 2018; Orey & Rosa, 2021). Integrating the ethnomathematics context into descriptive statistical problems will help students develop their statistical reasoning abilities.

The students' statistical reasoning abilities can also be developed through the support of Information and Communication Technology (ICT) in the application of descriptive statistics learning, in addition to integrating the ethnomathematics context in presenting statistical problems. The application of ICT in the mathematics learning process has a role as a medium that not only help students find solutions to the problems through visual displays but also have an important role in increasing students' exploration abilities related to a given mathematical problem. The process of finding and solving mathematical problems will be student-centered and this is very beneficial for students to understand what points they have not mastered from the materials provided. The application of ICT in descriptive statistics learning helps students explore the statistical data presented so that they can perform analysis and draw appropriate conclusions and understand the meaning of the statistical calculation process that has been carried out.

One of the ICT media that can be applied to help students explore statistical data is TinkerPlots. TinkerPlots is a software used to analyze exploratory data related to descriptive statistics that is carried out by students based on project learning. The TinkerPlots has dynamic visualization aspects that support and construct the way students visualize statistical models. According to Konold & Lehrer (2008) and Rubin et al. (2006) the objects that students make with this software and the outcomes of their excursions to arrange and study the outputs are a type of dynamic mathematical representation that prompts and supports their thinking. Additionally, this program assists students in developing and strengthening their abilities to translate statistical problems into TinkerPlots models, produce data using the models, and solve statistical problems using the data provided by the TinkerPlots models (Biehler et al., 2018; Rubin et al., 2006).

Several studies related to ICT as a learning medium have been conducted and published. Some of them use GeoGebra (Dynamic Software-2D) in learning mathematics (Bhagat & Chang, 2015; Günster & Weigand, 2020; Ramadhani & Narpila, 2018; Vahey et al., 2020); GeoGebra (Dynamic Software-3D) in learning mathematics (Velázquez & Méndez, 2021); Geometer's Sketchpad (Dynamic Software) in mathematics learning (Jungić et al., 2015; Mackrell, 2011; Soury-Lavergne & Maschietto, 2015); Desmos Mathematics Software (Hollebrands & Lee, 2020; Rocha, 2020); Cabri 3D in mathematics learning (Nurjanah et al., 2020); Autograph in mathematics learning(Bina et al., 2021; Moksin et al., 2018; Sari et al., 2020); and TinkerPlots in statistics learning (Aridor & Ben-Zvi, 2018a; Seloraji & Leong, 2016). Likewise, with studies that publish technology integration in various learning models. The integration of flipped classroom model with LMS-Google Classroom (Ramadhani, 2020); flipped classroom model with SIGIL-based E-Modul (Ramadhani & Fitri, 2020); Guided Discovery Learning model assisted by Autograph (Ramadhani, 2017); problem-based model Autograph-assisted learning (Ridha et al., 2018); and GeoGebra-assisted project-based learning models (Suherman et al., 2020).

However, until now, there has been no research that integrates the ethnomathematics approach with technology, one of which is TinkerPlots. The TinkerPlots application also has a novelty and uniqueness in learning mathematics because it is devoted to helping students in the process of critical thinking, statistical reasoning, and creative thinking in statistical problems. Previous research has not conducted specific research to examine how changes in students' abilities before and after being given the application of technology-based learning. The previous research only focused on seeing the considerable improvements that occurred before and after implementing technology-based learning. Therefore, this study provides a new perspective by analyzing changes in students' abilities, especially statistical reasoning abilities using the Stacking-Rasch Model Measurement analysis technique.

The application of informal problems and TinkerPlots in descriptive statistics learning based on the ethnomathematics approach is expected to assist students in developing their statistical reasoning abilities. In this study, the researchers focus on measuring the changes in statistical reasoning abilities of students who have received descriptive statistics learning using an ethnomathematics approach assisted by TinkerPlots. This study analyzes the pattern of changes in the statistical reasoning ability of each student and the difference of changes in statistical reasoning abilities of students who are taught by applying the ethnomathematics approach assisted by TinkerPlots and those who are only taught using ordinary learning. Based on the description of the findings, facts, and solutions offered in this study, the problems raised in this study are: (1) Is there any significant change in the students' statistical reasoning abilities after being given a lesson in the experimental group and the control group?; and (2) In comparison, how do students' statistical reasoning abilities change through the intervention of the TinkerPlots-assisted ethnomathematics approach and students who receive ordinary learning?

B. METHODS

This study aims to compare the amount to which students' statistical reasoning abilities can alter following the intervention using a quantitative method with a quasi-experimental design and a pre-test, post-test control group design. For five meetings, the researchers facilitate learning by administering tests and collecting data on the outcomes of treatments and assessments. The study sampled 72 12th-grade students from one of the city's high schools in Gunungsitoli, Nias islands, Sumatra Utara province. Following summarizes the demographic features of the research sample, as shown in Table 1.

Table 1. Demographic Features of Research Sample

Indicator		The number of students	Percentage
Gender	Male	41	56.94%
	Female	31	43.06%
Age	15-16 years	10	13.88%
	17-18 years	58	80.56%
	> 18 years	4	5.56%

The sample in this study is divided into two groups; the experimental group (a group of students who received an intervention using an ethnomathematical approach assisted by TinkerPlots) and the control group (a group of students who received an intervention using ordinary learning). The demographic features based on learning groups are displayed in Table 2.

Table 2. Demographic Features of Research Sample Based on Learning Group

Learning Group	Indicator		The number of students	Percentage
Experiment Group	Gender	Male	15	41.67%
		Female	21	58.33%
	Age	15-16 years	5	13.89%
		17-18 years	30	83.33%
		> 18 years	1	2.78%
Control Group	Gender	Male	17	47.22%
		Female	19	52.78%
	Age	15-16 years	5	13.89%
		17-18 years	28	77.78%
		> 18 years	3	8.33%

Based on Table 2, the number of students in the two learning groups is the same - 36 students. The changes are analyzed based on the pre-test and post-test given to each learning group. The results of the pre-test and post-test are then analyzed using the stacking analysis technique as part of the Rasch Measurement Model and analyzed using the WINSTEPS software.

Rasch Model Analysis is a type of measurement analysis that employs the Joint Maximum Likelihood Estimation (JMLE) equation to convert unstructured data to interval data (logit) (Soeharto, 2021). Rasch estimation is based on the interaction of item-person and probability estimation. The logit function can numerically express an item's interaction with a person or a student (log unit odd). The difficulty of the item and the person decide the chance of measurement concurrently, providing the item logit value (using each item's odd probability) and the person logit value (using the odd probability of each respondent). As a result, the probability is inversely proportional to the distance between two items (Boone, 2016; Boone et al., 2014). The Rasch analysis is used to overcome numerous challenges in this research by utilizing the Classical Test Theory (CTT). The Rasch analysis demonstrates that regardless of the item's difficulty level, the students' ability to measure remains constant. Still, the item's difficulty level remains constant irrespective of the students' knowledge (Ramadhani & Evans, 2022).

Rasch Model Measurement has developed various measurement models, ranging from straightforward dichotomous models to sophisticated tests. Stacking analysis is one of the data analyses performed in this study on the Rasch Model. Stacking analysis is a technique for vertically stacking pre-and post-test data (Combrinck et al., 2017). Each student has two rows of data, one for the pre-test and another for the post-test. The data set contains two instances of each student, although the applicable test item appears only once in the experimental and control groups. Due to the data location, the researcher can examine changes in individual students following the intervention. Each student is evaluated on the same test question, which enables comparisons between pre-and post-test ability. Students' abilities on each pre-post-test item may be compared since the data is analyzed in a single measurement, creating an item measurement for each student and one item for the entire class (Laliyo, 2021; Uzun & Öğretmen, 2021).

This study collected pre-test data prior to administering the intervention to each learning group. Additionally, post-test data collecting occurs following the delivery of the intervention. The pre- and post-test items are constructed identically. The pupils complete the problems on the answer sheet provided. While collecting students' test results, the measurement data before to and following the test are still in the form of ordinal data. The Rasch Partial Credit Model (RPCM) is used to transform ordinal data to interval data with the same logit scale, helped by the WINSTEPS 4.5.5 software. The process of data transformation occurs as a result of data calibration on the ability and difficulty level of items within the same interval.

The results of the pre-test and the post-test data analysis in logit values that have been analyzed by applying the stacking technique are then taken to test the significance of the changes that occur before and after the intervention. The significance test is carried out by applying the Man-Whitney U test and Wilcoxon test. The significance test using the Man-Whitney U test and Wilcoxon test can be done if the students' data of the pre-test and post-test logit values are homogeneous and normally distributed. The hypothesis in this study is that the students' statistical reasoning ability from pre-test to post-test in the group that receives the TinkerPlots-assisted ethnomathematics approach intervention is changed positively and significantly.

C. RESULT AND DISCUSSION

1. Rasch Analysis Properties of Instrument

The test instrument applied in this research is a statistical reasoning ability test in five essay questions. The test instrument applied in this research is initially tested by applying a unidimensionality test, rating scale test, reliability test on both person and item, person separation index test, and validity test using Rasch Model Measurement analysis. The instrument is previously tested on 120 high school students who are not the sample in this study. The unidimensionality test is carried out to optimize the measurements constructed so that the resulting information is more focused on the measured attribute (in this case, the ability of statistical reasoning). The instrument's unidimensionality is critical in determining if the designed instrument can accurately measure what it is intended to measure (Laliyo, 2021). In the unidimensionality test, the residuals are subjected to Principal Component Analysis, which determines the extent to which the change in the instruments accurately measures what should be assessed (Sumintono & Widhiarso, 2015). Following summarizes the findings of the unidimensionality test on the statistical reasoning ability test instrument, as shown in Table 3.

Table 3. Standarized Residual Variance in Eigenvalue Units (Unidimensionality Measure)

	Eigenvalue	**Observed**		**Expected**
Total observed variance in raw form	12.5121	100%		100%
Measures that account for raw variation	7.5121	60.0%		58.8%
Person explains the raw variance.	4.0936	32.7%		32.1%
The variance in raw values explained by items	3.4185	27.3%		26.8%
Variance that is inexplicable in its raw form (total)	5.0000	40.0%	100.0%	41.2%
Variation in the first contrast that is unexplained	1.9249	15.4%	38.5%	
Variation in the second contrast that is unexplained	1.1728	9.4%	23.5%	
Unaccounted-for variation in the third contrast	1.0128	8.1%	20.3%	
Unaccounted-for variation in the fourth contrast	.8815	7.0%	17.6%	
Unaccounted-for variation in the fifth contrast	.0087	.1%	.2%	

The results presented in Table 3 show that the test instrument for the reasoning ability relatively has a good measure of unidimensionality. It shows that the raw variance measurement result is 60.0%. This value indicates that the unidimensionality requirement is excellent since the minimum value for this requirement is 20% and it is considered remarkable if it reaches more than 60%. Another value shown in Table 1 is the value of the raw unexplained variance (the unexplainable variance by the instrument), which ideally does not exceed 15%. Moreover, the results obtained in this study show a value of 5%. It means that the reasoning ability test instrument can effectively measure the students' statistical reasoning ability as it shows the value below 10%.

The next stage is to test the rating scale. The rating scale is carried out to test whether the assessment criteria or the rating of students' reasoning abilities can be used or not. The Rasch model can provide conclusions for the rating assumptions applied in the statistical ability reasoning instrument, which can be observed from the Observed Average and Andrich Threshold. The results of the analysis as shown in Table 4.

Table 4. Summary of Category Structure

Category Label	Score	Observed		Observed Average	Sample Expect	INFIT MNSQ	OUTFIT MNSQ	Andrich Threshold	Category Measure	Rating Scale
		Count	%							
1	1	4	1	− .99	−2.42	1.95	1.31	NONE	(−6.38)	1
2	2	148	25	−.80	−.88	1.03	1.03	−5.28	−2.91	2
3	3	227	38	.48	.67	1.01	1.09	−.53	.60	3
4	4	162	27	2.39	2.26	.80	.76	1.79	2.93	4
5	5	59	10	3.72	3.68	.95	.94	4.01	(5.18)	5

OBSERVED AVERAGE is mean of measure in category. It is not a parameter estimate.

Table 4 shows that the observed average starts from logit −.99 for the first rank (very poor statistical reasoning ability), logit −.80 for the second rank (poor statistical reasoning ability), logit .48 for the third rank (fair statistical reasoning ability), logit 2.39 for the fourth rank (good statistical reasoning ability), and logit 3.72 for the fifth rank (excellent statistical reasoning ability)(Laliyo, 2021).

Further testing is carried out to observe the reliability of both person and item. The reliability testing on the test instrument aims to find out how far the measurement generates consistent information in revealing latent traits or the nature of the measured unidimensional variables ur (Sumintono & Widhiarso, 2015). The results of the analysis are presented in summary statistics as shown in Table 5.

Table 5. Reliability of Pearson and Item

	Person (120)	Item (5)
Reliability	.74	.98
Separation	1.69	6.53
Measure (SD)	.99 (.15)	.00 (.52)
INFIT MNSQ	.97	.99
INFIT ZSTD	−.06	−.21
OUTFIT MNSQ	.97	.97
OUTFIT ZSTD	−.04	−.32
KR (20) = .76		

Table 5 shows that the person reliability value is .74 which is equivalent to the person separation index value of 1.69. It also shows that the consistency of students' responses to the statistical reasoning ability test instrument is quite good. The .76 value of the Cronbach Alpha Coefficient (KR-20) indicates that there is good interaction between students and the statistical reasoning ability test instrument. This result indicates a strong correlation between students' responses and items, where the students' knowledge tends not to be fragmented so that it can be measured. Based on this result, it can be stated that the statistical reasoning ability test instrument is responsive and reliable enough to clearly distinguish students' statistical reasoning abilities. In addition, the acquisition of the item separation index value of 6.53 is quite high, equivalent to the item reliability value of .98. These results indicate that the consistency of the items is remarkable, or the items can be stated to meet the unidimensionality requirements. These findings also point out that the items are able to competently define the measured variable. This conclusion is confirmed by the acquisition of INFIT and OUTFIT item values, most of which were within the acceptable range for the essay test (Bond & Fox, 2015).

Table 5 also includes statistics on the person separation index. The person separation index is used to determine the degree to which the statistical reasoning ability test can discriminate between students' statistical reasoning abilities when presented with descriptive statistics content. The higher the value of the person separation index, the more likely it is that students will respond to items rationally. The item separation index indicates the degree to which things are widely distributed when classifying simple and difficult items. The more goods that are available, the better and more relevant they are (Boone et al., 2014). The person separation index (1.69) and item separation index(6.53) in this study reflect a reasonably wide range of statistical reasoning ability test instruments on the basis of person and item. This measure validates the statistical reasoning ability test as a valid and reliable instrument for assessing students' statistical reasoning abilities when presented with descriptive statistics content.

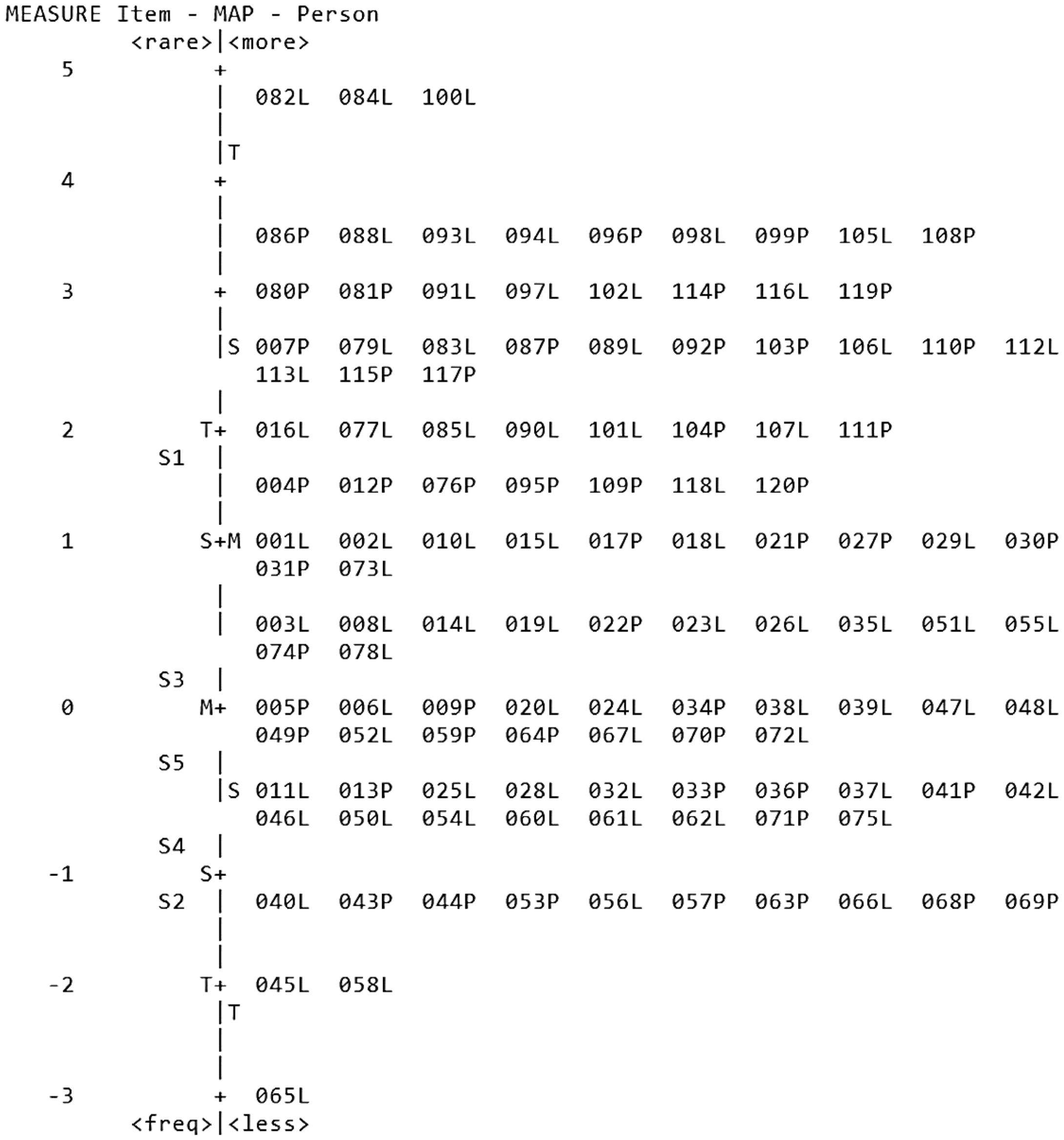
The final test in the research instrument test is the validity test. The validity analysis is intended to test whether the statistical reasoning ability test instrument measures what it is intended to. The statistical reasoning ability test instrument is considered to have good construct validity if it is able to measure changes in the level of students' statistical reasoning ability (Chan et al., 2014, 2016). The first step of construct validity is to ensure that all items of statistical reasoning ability match the Rasch Model. An item is said to be a misfit if the obtained measurement results do not match the following criteria, there are: (i) the accepted value of Outfit Mean Square (MNSQ) is. 5 to 1.5 ; (ii) The accepted value of Outfit Z-Standard (ZSTD) is -2.0 to $+2.0$; and (iii) The value of Point Measure Correlation (Pt. Mean Corr) is 0.4 to 0.85.

Summary data are used in the instrument reliability testing. It is worth noting that the sample size has a significant effect on the ZSTD value. When a large sample size is employed, the ZSTD value is always more than three ($\mathrm{ZSTD} > 3$). Thus, some experts advise against utilizing the ZSTD value in studies with a large sample size ($N > 500$) (Boone et al., 2014; Sumintono & Widhiarso, 2015). Mean Square Residual (MNSQ) logit value indicates the magnitude of the item misfit effect. Outfit MNSQ and Infit MNSQ are also presented. Following summarizes the findings from the item statistics analysis, as shown in Table 6.

Table 6. Item Statistics: Misfit Order

Item	Measure	Outfit MNSQ	Outfit ZSTD	PTMEA CORR
S1	1.80	1.27	1.71	.39
S2	.35	.97	−.17	.83
S3	−.20	.84	−1.37	.82
S4	−.79	.83	−1.44	.81
S5	−1.16	.95	−.34	.60
Mean	.00	.97	−.3	
P.SD	1.04	.16	1.1	

The results of the item statistics analysis based on Table 6 show that all of the items have met the criteria for the Rasch model validity (S1: Item test number 1; S2: Item test number 2; S3: Item test number 3; S4: Item test number 4; and S5: Item . The presentation of item statistics is also displayed in the Wright map presented in Figure 1.



```
MEASURE Item - MAP - Person
         <rare>|<more>
    5          +
               |  082L  084L  100L
               |
               |T
    4          +
               |
               |  086P  088L  093L  094L  096P  098L  099P  105L  108P
               |
    3          +  080P  081P  091L  097L  102L  114P  116L  119P
               |
               |S 007P  079L  083L  087P  089L  092P  103P  106L  110P  112L
                  113L  115P  117P
               |
    2         T+  016L  077L  085L  090L  101L  104P  107L  111P
           S1  |
               |  004P  012P  076P  095P  109P  118L  120P
               |
    1         S+M 001L  002L  010L  015L  017P  018L  021P  027P  029L  030P
                  031P  073L
               |
               |  003L  008L  014L  019L  022P  023L  026L  035L  051L  055L
                  074P  078L
           S3  |
    0         M+  005P  006L  009P  020L  024L  034P  038L  039L  047L  048L
                  049P  052L  059P  064P  067L  070P  072L
           S5  |
               |S 011L  013P  025L  028L  032L  033P  036P  037L  041P  042L
                  046L  050L  054L  060L  061L  062L  071P  075L
           S4  |
   -1         S+
           S2  |  040L  043P  044P  053P  056L  057P  063P  066L  068P  069P
               |
               |
   -2         T+  045L  058L
               |T
               |
               |
   -3          +  065L
         <freq>|<less>
```

Figure 1. Wright Map Person ($N = 120$) and Item ($N = 5$)

Based on Figure 1, all of the items in the statistical reasoning ability instrument cover most of the students' statistical reasoning abilities. There are only three students below the 2nd item (S2) difficulty level. Otherwise, most of the students are above the 1st (S1) and 3rd item (S3) where those items are the highest difficulty level. In addition, there are no items that are not in accordance with the student's ability. Overall, the difficulty of the items aligned with the construct.

2. The Changes in Students' Statistical Reasoning Skills of Descriptive Statistics

The statistical reasoning test instrument that has proceeded through the instrument testing process can further be applied to measure changes in the statistical reasoning ability of the research sample. Based on the results of instrument testing, it is found that all of the instruments designed had met the criteria for instrument testing, which is five questions related to statistical reasoning abilities. The statistical reasoning test instrument is then provided to the research sample; the students in the experimental group (receiving a learning intervention using TinkerPlots-assisted ethnomathematics approach) and students in the control group (receiving a learning intervention using ordinary learning). The data collected during the intervention are analyzed using the Rasch Measurement Model and the Stacking analysis technique. The test results will provide a logit value that will be used to compare students' statistical reasoning abilities prior to and following the intervention. The logit value is calculated using the raw scores on tests administered to college students. Thus, the stacking technique used in the Rasch Model Measurement does not take into account the students' aggregate final score, but rather the results for each essay test taken by each student, in accordance with the application of the Item Response Theory (IRT). The following Wright Map illustrates the logit value changes in students' pre- and post-test scores in both experimental and control groups (Figure 2 and Figure 3).

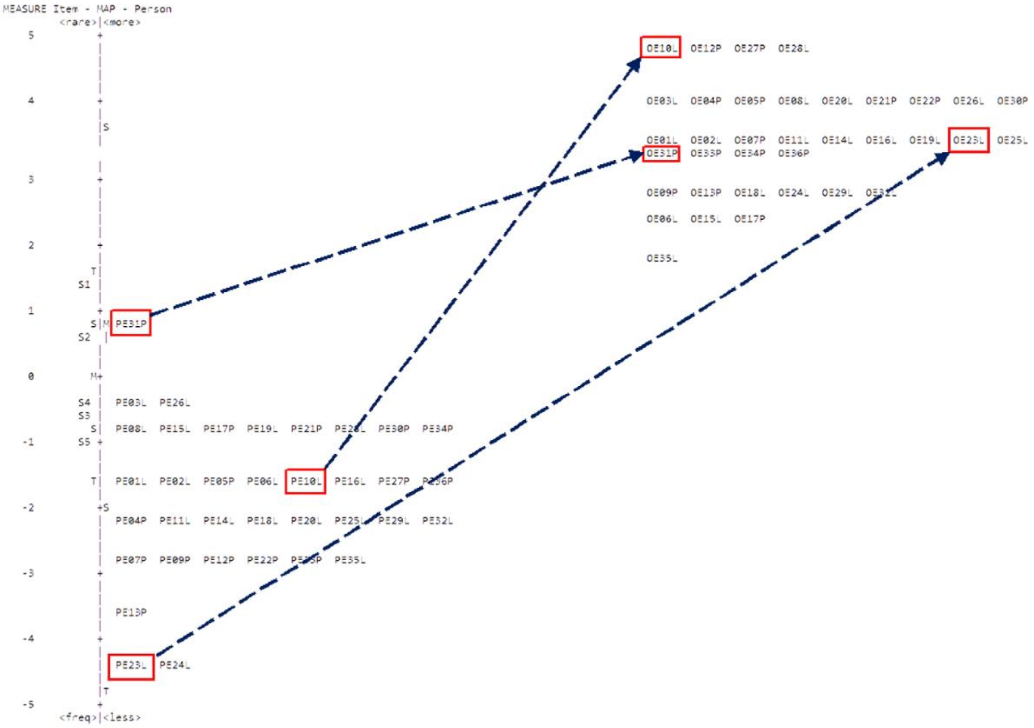


Figure 2. Wright Map of Experiment Class

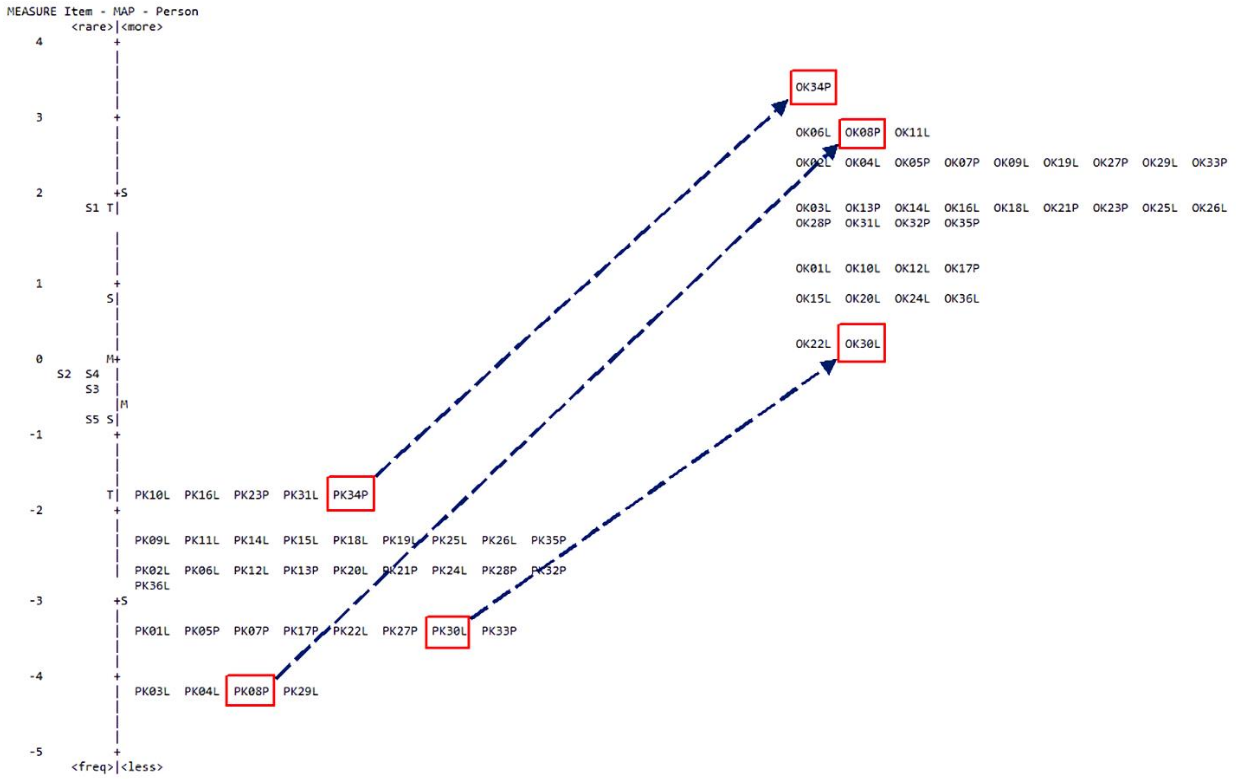


Figure 3. Wright Map of Control Class

Based on Figure 2 and Figure 3, it is found that the average change in logit value (pre-test and post-test) in the two groups is distinct. The average change in the experimental group logit value is higher than the average change in the control group logit value ($5.335 > 4.664$). Figure 2 shows that all students in the experimental group are higher than $+2.00$ and can solve problems in the difficult category - the first item (S1) and the second item (S2). It appears that student code PE23L (the student with code 23, male) is the student who has the lowest logit value during the pre-test with a value below -4.00 (the student's logit value is -4.42). However, it can be seen that during the post-test, the student code OE23L has a logit value change above $+3.00$ (OE23L logit value is $+3.45$). The same circumstance is also observed in the other student - the student code PE10L (student code 10, male) with logit values below -1.00 (the student's logit value is -1.50). However, in the post-test, the student code OE10L has the logit value change above $+4.00$ (the student's logit value is $+4.78$) and is the student having the highest post-test score in the experimental group. The student with the highest logit value (logit value $+0.79$) in the pre-test, the student code PE31P (students code 31, female), also experiences a change in the logit value in the post-test, which is above $+3.00$ (the student's logit value is $+3.45$) but does not achieve the optimal change in value as experienced by PE23L code student.

Similar circumstances are also experienced by the students in the control group. Based on Figure 3, it appears that all students in the control group are above 0.00 logit value and can solve the easy category questions - the fifth item (S5). However, they still have difficulty solving the difficult category questions - the first item (S1). It appears that the student code PK08P (student code 08, female) is the student with the lowest logit value at the pre-test, which is below −4.00 (the student's logit value is −4.11). However, there is a great change in the post-test conditions, where student code OK08P obtains the logit value change above 0.00 (the student's logit value is +2.85). The changes by 08P are the optimal changes experienced by students from pre-test to post-test. The change in the pre-test and post-test scores that occurs in 08P is 6.96. There are other findings in Figure 3, where the student code PK30L (student code 30, male) also experiences a remarkable change in logit value in the pre-test with a logit value below −3.00 (the student's logit value is −3.48). In the post-test, student code OK30L experiences a logit value change below +1.00 (the student's logit value is +0.12). The other condition is found in student code PK34P, one of the students with the highest logit value in the pre-test, who has a logit value of −1.00 (the student's logit value is−1.75) and in the post-test, the student code OK34P experiences optimal and constant changes to become the student with the highest logit score – above +3.00 (the student's logit value is +3.38). Moreover, the student code K34P shows that student with good initial statistical reasoning abilities can take advantage of these abilities to improve their abilities further.

The student code E23L in the experimental group experiences changes in the optimal logit value from pre-test to post-test. The optimal change in logit value is caused by the intervention carried out in the experimental group-TinkerPlots-assisted ethnomathematics approach. The intervention given to the students in the experimental group provides a new learning experience and gives the students more flexibility in exploring the given problem. The students also seem to enjoy participating in learning activities using TinkerPlots. Andre et al. (2019) and Fitzallen (2013) agree that TinkerPlots obviously plays a major role in increasing students' motivation and interest in learning so that it has an impact on students' curiosity about solving problems. The curiosity also gives a positive response to students in improving the statistical reasoning experienced during solving informal problems based on the Nias cultural context. TinkerPlots is also very helpful for students to ensure that the statistical reasoning process carried out previously strengthens the hypotheses obtained by students in solving informal problems in Nias cultural context. Bakker et al. (2006) and Fitzallen & Watson (2010) conclude that TinkerPlots also assist students in supporting their further conceptual development.

The condition experienced by student code K34P indicates that the student already has good statistical reasoning ability and the student does not experience difficulties in improving. These results conclude that the intervention given to students in the control group is not the only factor that plays a role in increasing students' statistical reasoning abilities, but the student's initial ability factor can also be a factor affecting the improvement obtained (ten Braak et al., 2022; Wyse & Mapuranga, 2009). The interventions carried out through ordinary learning do not provide new learning experiences to students, because the learning provided is no different from the regular learning. This condition makes students less enthusiastic and less motivated to participate in learning activities, but it cannot be denied that intervention through ordinary learning still provides changes in students' statistical reasoning abilities from pre-test to post-test.

3. The Difference in Students' Statistical Reasoning Skills of the Descriptive Statistics

The students' statistical reasoning abilities are evaluated using a significance test of differences on descriptive statistics material utilizing pre- and post-test data collected from the

Person Statistics: Entry Order Rasch Model test. Prior to the test, the homogeneity and data normality of the pre- and post-test logit data (produced from the Person Statistics: Entry Order Rasch Model test) are determined. The results of the normality and homogeneity tests on the pre- and post-test logit data in the groups indicate that the data are not homogeneous or normally distributed. To determine the difference in ability between the experimental and control groups on the pre- and post-test, a non-parametric test, the Mann Whitney U Test, is used. The Wilcoxon signed-rank test is also used to examine the differences between non-normally distributed data pairs. The Mann-Whitney U and Wilcoxon tests are employed in this study to assess if there is a statistically significant difference in statistical reasoning skills within and between experimental and control groups. The analysis's findings are presented in Tables 7 and 8.

Table 7. Mann-Whitney U Test Result Based on Students' Pre-Test and Post-Test of Statistical Reasoning Ability in Experimental and Control Group ($p < .05$)

Test	Experimental Group ($\mathbf{N = 36}$)	Control Group ($\mathbf{N = 36}$)	U	p
Pre-Test	-1.8108	3.5231	228.000	$.000$
Post-Test	-2.8558	1.8083	31.000	$.000$

Table 8. Wilcoxon Test Result Based on Students' Pre-Test and Post-Test of Statistical Reasoning Ability in Experimental and Control Group ($p < .05$)

Group	Pre-Test	Post-Test	Z	p
Experimental	-1.8108	3.5231	-3.866^b	$.000$
Control	-2.8558	1.8083	-5.188^b	$.000$
[b] Based on positive ranks				

Based on Table 7, the results of the Mann-Whitney U test show that there is a statistically significant difference between the results of the pre-test ($\mathrm{U} = 228.000, p < .05$) and the post-test ($\mathrm{U} = 31.000, p < .05$) in the experimental group and control group. The Wilcoxon test results in Table 6 also show that the experimental group and control group have statistically significant reasoning abilities based on the pre-test result ($\mathrm{Z} = -3.866^b, p < .05$) and post-test results -test ($\mathrm{Z} = -5.188^b, p < .05$). Based on these results, it can be concluded that the statistical reasoning ability of the students in the post-test is much higher than in the pre-test for both groups. However, the improvement in statistical reasoning ability is better in the experimental group compared to the control group. These results mean that the research hypothesis is valid as the students' statistical reasoning ability from pre-test to post-test in the TinkerPlots-assisted ethnomathematics approach group changes positively and significantly. The results also conclude that students who receive learning interventions using the TinkerPlots-assisted ethnomathematics approach perform better than students experiencing ordinary learning.

Based on the validation of the hypothesis, the intervention of the TinkerPlots-assisted ethnomathematics approach is effective to improve the statistical reasoning ability of high school students. This finding also supports research conducted by Seloraji & Leong (2016), Aridor & Ben-Zvi (2018b), and English & Watson (2015) that TinkerPlots helps students in understanding and reasoning statistical data presented visually, and solving statistical problems. Risdiyanti & Prahmana (2020), Rosa & Orey (2017), and Putra (2018) also confirm that the application of ethnomathematics-based informal problems helps students understand mathematical problems through their traditions and culture. The integration of ethnomathematics also provides new learning experiences for improving students' behavior and personal factors so that they have a significant impact on improving students' mathematical abilities (Ramadhani et al., 2021).

D. CONCLUSION AND SUGGESTIONS

The stacking analysis on logit data shows that the statistical reasoning ability of students experiences significant changes before and after the intervention is given. It occurs in both students receiving learning interventions using the TinkerPlots-assisted ethnomathematics approach and ordinary learning. This finding provides an interpretation that the learning intervention applied in the experimental group and in the control group is proven to be able to provide changes to the statistical reasoning ability of high school students. Interventions given to students are effective in helping students improve their statistical reasoning abilities. The stacking analysis also shows that changes in abilities experienced by students can be measured using data during pre-test and post-test without using the final data. The results of the stacking analysis provide new findings where students with any initial ability (in the pre-test condition) have the same opportunity to improve the ability (in the post-test) so that every intervention carried out still has an impact on the students' mathematical abilities.

The findings obtained from the stacking analysis also support measuring the differences in students' statistical reasoning abilities in the pre-test and post-test in both groups using the difference test. The results of the difference test show that the students have higher reasoning abilities in the post-test than in the pre-test. This result is indeed obtained from the intervention given to both groups. The intervention provided has a significant impact and results in changing the students' abilities during the post-test. The results of the difference test also show that students who receive a learning intervention by the TinkerPlots-assisted ethnomathematics approach are better than students who receive ordinary learning. This finding provides a meaningful interpretation where the application of the TinkerPlots-assisted ethnomathematics approach is effective in helping students improve their statistical reasoning ability. This study also proves that the application of the TinkerPlots application in mathematics learning has encouraged the development of students' statistical reasoning abilities. The stacking-Rasch Model technique manifests an effective way of examining data collected multiple times. Researchers who want to conduct research in developing the other mathematical abilities can employ this finding to obtain more significant analysis results on changes in student abilities during pre-test and post-test

ACKNOWLEDGEMENT

We would like to express our gratitude to Universitas Negeri Medan and Universitas Potensi Utama for providing facilities and opportunities to complete and collaborate on this project.

REFERENCES

Andre, M., Lavicza, Z., & Prodromou, T. (2019). Formalizing students' informal statistical reasoning on real data: Using Gapminder to follow the cycle of inquiry and visual analyses. *Eleventh Congress of the European Society for Research in Mathematics Education (CERME11)*.

Aridor, K., & Ben-Zvi, D. (2018a). Statistical modeling to promote students' aggregate reasoning with sample and sampling. *ZDM* , *50*(7), 1165–1181. https://doi.org/10.1007/S11858-018-0994-5

Aridor, K., & Ben-Zvi, D. (2018b). Statistical modeling to promote students' aggregate reasoning with sample and sampling. *ZDM - Mathematics Education*, *50*(7), 1165–1181. https://doi.org/10.1007/s11858-018-0994-5

Bakker, A., Derry, J., & Konold, C. (2006). Using Technology to support diagrammatic reasoning about center and variation. In A. Rossman & B. Chance (Eds.), *Working Cooperatively in Statistics Education. Proceedings of the Seventh International Conference on Teaching Statistics, Salvador, Brazil. [CDROM]*. International Statistical Institute.

Bhagat, K. K., & Chang, C.-Y. (2015). Incorporating GeoGebra into Geometry learning-A lesson from India. *Eurasia Journal of Mathematics, Science & Technology Education*, *11*(1), 77–86.

Biehler, R., Frischemeier, D., & Podworny, S. (2018). Elementary preservice teachers' reasoning about statistical modeling in a civic statistics context. *ZDM* , *50*(7), 1237–1251. https://doi.org/10.1007/S11858-018-1001-X

Bina, N. S., Fitri, Y., Sihotang, S. F., & Saragih, R. M. B. (2021). Use of autograph learning media to improve mathematic communication skills. *Advances in Social Science, Education and Humanities Research*, 86–91. https://doi.org/10.2991/assehr.k.210413.021

Bond, T. G., & Fox, C. M. (2015). *Applying in the Rasch Model, Fundamental Measurament in the Human Sciences* (3rd editio). Routledge.

Boone, W. J. (2016). Rasch analysis for instrument development: Why,when,and how? *CBE Life Sciences Education*, *15*(4). https://doi.org/10.1187/cbe.16-04-0148

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

Chan, S. W., Ismail, Z., & Sumintono, B. (2014). A Rasch Model Analysis on Secondary Students' Statistical Reasoning Ability in Descriptive Statistics. *Procedia - Social and Behavioral Sciences*, *129*, 133–139. https://doi.org/10.1016/j.sbspro.2014.03.658

Chan, S. W., Ismail, Z., & Sumintono, B. (2016). A Framework for Assessing High School Students' Statistical Reasoning. *PLoS ONE*, *11*(11), 1–32. https://doi.org/10.1371/journal.pone.0163846

Ching, B. H. H., Kong, K. H. C., Wu, H. X., & Chen, T. T. (2020). Examining the reciprocal relations of mathematics anxiety to quantitative reasoning and number knowledge in Chinese children. *Contemporary Educational Psychology*, *63*, 101919. https://doi.org/10.1016/j.cedpsych.2020.101919

Combrinck, C., Scherman, V., & Maree, D. (2017). Evaluating anchor items and reframing assessment results through a practical application of the Rasch Measurement Model. *South African Journal of Psychology*, *47*(3), 316–329. https://doi.org/10.1177/0081246316683569

D'Ambrosio, U. (2018). The program ethnomathematics: Cognitive, anthropological, historic and socio-cultural bases. *PNA*, *12*(4), 229–247. https://doi.org/10.30827/pna.v12i4.7851

delMas, R. C. (2002). Statistical Literacy, Reasoning, and Learning: A Commentary. *Journal of Statistics Education*, *10*(3). https://doi.org/10.1080/10691898.2002.11910679

English, Lyn D., & Watson, J. M. (2015). Exploring variation in measurement as a foundation for statistical thinking in the elementary school. *International Journal of STEM Education*, *2*(1). https://doi.org/10.1186/s40594-015-0016-x

Fitzallen, N. E. (2013). Characterising Students' Interaction with TinkerPlots. *Technology Innovations in Statistics Education*, *7*(1). http://www.escholarship.org/uc/item/1074n1dp

Fitzallen, N., & Watson, J. (2010). Developing Statistical Reasoning Facilitated By TinkerPlots. *Proceedings of the Eighth International Conference on Teaching Statistics (ICOTS8)*. www.stat.auckland.ac.nz/~iase/publications.php

Garfield, J. (2002). The Challenge of Developing Statistical Reasoning. *Journal of Statistics Education*, *10*(3). https://doi.org/10.1080/10691898.2002.11910676

Günster, S. M., & Weigand, H.-G. (2020). Designing digital technology tasks for the development of functional thinking. *ZDM* , *52*(7), 1259–1274. https://doi.org/10.1007/S11858-020-01179-1

Gürbüz, R., & Erdem, E. (2016). Relationship between mental computation and mathematical reasoning. *Cogent Education*, *3*(1), 1212683. https://doi.org/10.1080/2331186X.2016.1212683

Hollebrands, K. F., & Lee, H. S. (2020). Effective design of massive open online courses for mathematics teachers to support their professional learning. *ZDM* , *52*(5), 859–875. https://doi.org/10.1007/S11858-020-01142-0

Hughes, E. M., Riccomini, P. J., & Lee, J. Y. (2020). Investigating written expressions of mathematical reasoning for students with learning disabilities. *Journal of Mathematical Behavior*, *58*, 100775. https://doi.org/10.1016/j.jmathb.2020.100775

Im, S. H., & Jitendra, A. K. (2020). Analysis of proportional reasoning and misconceptions among students with mathematical learning disabilities. *Journal of Mathematical Behavior*, *57*, 100753. https://doi.org/10.1016/j.jmathb.2019.100753

Jungić, V., Kaur, H., Mulholland, J., & Xin, C. (2015). On flipping the classroom in large first year calculus courses. *International Journal of Mathematical Education in Science and Technology*, *46*(4), 508–520. https://doi.org/10.1080/0020739X.2014.990529

Justice, N., Zieffler, A., Huberty, M. D., & delMas, R. (2018). Every rose has its thorn: secondary teachers' reasoning about statistical models. *ZDM - Mathematics Education*, *50*(7), 1253–1265. https://doi.org/10.1007/s11858-018-0953-1

Kilpatrick, J., Martin, W. G., Schifter, D., & National Council of Teachers of Mathematics. (2009). *A research companion to principles and standards for school mathematics*. National Council of Teachers of Mathematics.

Konold, C., & Lehrer, R. (2008). Technology and mathematics education: An essay in honor of Jim Kaput. In L. D. English (Ed.), *Handbook of international research in mathematics education* (2nd editio, pp. 49–72). Taylor & Francis.

Laliyo, L. A. R. (2021). *Mendiagnosis Sifat Perubahan Konseptual Siswa: Penerapan Teknik Analisis Stacking dan Racking Rasch Model*. Deepublish.

Lepak, J. R., Wernet, J. L. W., & Ayieko, R. A. (2018). Capturing and characterizing students' strategic algebraic reasoning through cognitively demanding tasks with focus on representations. *Journal of Mathematical Behavior*, *50*, 57–73. https://doi.org/10.1016/j.jmathb.2018.01.003

Mackrell, K. (2011). Design decisions in interactive geometry software. *ZDM* , *43*(3), 373–387. https://doi.org/10.1007/S11858-011-0327-4

Mahdayani, R. (2016). Analisis kesulitan siswa dalam pemecahan masalah matematika pada materi aritmetika, aljabar, statistika. dan geometri. *PENDAS MAHAKAM: Jurnal Pendidikan Dan Pembelajaran Sekolah Dasar*, *1*(1), 86–98. https://jurnal.fkip-uwgm.ac.id/index.php/pendasmahakam/article/view/39

Moksin, A. I., Shahrill, M., Anshari, M., Huda, M., & Tengah, K. A. (2018). The Learning of integration in calculus using the autograph technology. *Advanced Science Letters*, *24*(1), 550–552. https://doi.org/10.1166/asl.2018.12067

Mumu, J., & Tanujaya, B. (2019). Measure Reasoning Skill of Mathematics Students. *International Journal of Higher Education*, *8*(6). https://doi.org/10.5430/ijhe.v8n6p85

Nuralam, & Gadeng, T. (2018). Kualitas Hasil Belajar Statistika Siswa MTs yang dibelajarkan dengan Pendekatan Problem Posing. *Suska Journal of Mathematics Education*, *4*(2), 88–99.

Nurjanah, Latif, B., Yuliardi, R., & Tamur, M. (2020). Computer-assisted learning using the Cabri 3D for improving spatial ability and self- regulated learning. *Heliyon*, *6*(11), e05536. https://doi.org/10.1016/J.HELIYON.2020.E05536

Orey, D. C., & Rosa, M. (2021). Ethnomodelling as a glocalization process of mathematical practices through cultural dynamism. *Mathematics Enthusiast*, *18*(3), 439–468.

Pusat Penilaian Pendidikan Kementerian Pendidikan dan Kebudayaan. (2022a). *Laporan Hasil Ujian Nasional-Penguasaan Materi Ujian Nasional*. Https://Hasilun.Puspendik.Kemdikbud.Go.Id/#2019!Sma!Daya_serap!07&08&999!A&03&1&T&1&unbk!3!&.

Pusat Penilaian Pendidikan Kementerian Pendidikan dan Kebudayaan. (2022b). *LAPORAN HASIL UJIAN NASIONAL 2015-2019*. Https://Hasilun.Puspendik.Kemdikbud.Go.Id/#2015-2016-2017-2018-2019!Sma!Capaian_nasional!07&99&999!T&T&1&N&1&unbk!1!&.

Putra, M. (2018). How ethnomathematics can bridge informal and formal mathematics in mathematics learning process at school: A framework. In *For the Learning of Mathematics* (Vol. 38, Issue 3). https://www.scopus.com/inward/record.uri?partnerID=HzOxMe3b&scp=85073814935&origin=inward

Ramadhani, R. (2017). Peningkatan kemampuan pemahaman konsep dan kemampuan pemecahan masalah matematika siswa SMA melalui guided discovery learning berbantuan autograph. *Jurnal Penelitian Dan Pembelajaran Matematika*, *10*(2), 72–81. https://doi.org/10.30870/jppm.v10i2.2032

Ramadhani, R. (2020). The effect of flipp-blended learning model using LMS-google classroom in mathematics learning. *Proceedings of the 1st International Conference of Global Education and Society Science, ICOGESS 2019*. https://doi.org/10.4108/EAI.14-3-2019.2292028

Ramadhani, R., & Evans, B. (2022). Measuring students' statistical reasoning abilities using flipped classroom model with SPSS and STATCAL. *Journal of Honai Math*, *5*(1), 1–14. https://doi.org/10.30862/jhm.v5i1.246

Ramadhani, R., & Fitri, Y. (2020). A Project-based learning into flipped classroom for ePUB3 electronic mathematics learning module (eMLM)-based on course design and implementation. *Universal Journal of Educational Research*, *8*(7), 3119–3135. https://doi.org/10.13189/ujer.2020.080740

Ramadhani, R., & Narpila, S. D. (2018). Problem based learning method with geogebra in mathematical learning. *International Journal of Engineering and Technology(UAE)*, *7*(3.2 Special Issue 2).

Ramadhani, R., Syahputra, E., & Simamora, E. (2021). Ethno-Flipped Classroom Model: Sebuah Rekomendasi Model Pembelajaran Matematika di Masa New Normal. *AXIOM: Jurnal Pendidikan Dan Matematika*, *10*(2), 221–240. https://doi.org/10.30821/axiom.v10i2.10331

Ridha, M., Johar, R., & Mailizar. (2018). The effectiveness of an autograph-assisted problem based learning model for improving high school students' learning outcome on graphics functions. *IOP Conf. Series: Journal of Physics: Conf. Series 1088*, 012025. https://doi.org/10.1088/1742-6596/1088/1/012025

Risdiyanti, I., & Prahmana, R. C. I. (2020). *Ethnomathematics (Teori dan Implementasinya: Suatu Pengantar)*. UAD Press.

Rocha, H. (2020). Using tasks to develop pre-service teachers' knowledge for teaching mathematics with digital technology. *ZDM* , *52*(7), 1381–1396. https://doi.org/10.1007/S11858-020-01195-1

Rosa, M., & Orey, D. C. (2017). STEM education in the Brazilian context: An ethnomathematical perspective. In *STEM Education in the Junior Secondary: The State of Play* (pp. 221–247). Springer, Singapore. https://doi.org/10.1007/978-981-10-5448-8_11

Rubin, A., Hammerman, J., & Konold, C. (2006). Exploring informal inference with interactive visualisation software. In A. Rossman & B. Chance (Eds.), *Working Cooperatively in Statistics Education. Proceedings of the Seventh International Conference on Teaching Statistics*. International Statistical Institute. http://iaseweb.org/documents/papers/icots7/2D3_RUBI.pdf

Sari, L., Syahputra, E., & Surya, E. (2020). Development of Autograph-Based Learning Tools to Improve Mathematical Communication Skills Students in vocational high school. *International Journal of Multicultural and Multireligious Understanding*, *7*(8), 326. https://doi.org/10.18415/ijmmu.v7i8.1880

Seloraji, P., & Leong, K. E. (2016). Impact of using Tinkerplots on statistical reasoning. *Proceedings the 21st Asian Conference on Technology in Mathematics (Pattaya, Thailand)*.

Setiawan, E. P., & Sukoco, H. (2021). Exploring first year university students' statistical literacy: A case on describing and visualizing data. *Journal on Mathematics Education*, *12*(3), 427–448. https://doi.org/10.22342/jme.12.3.13202.427-448

Sharma, S. (2017). Definitions and models of statistical literacy: a literature review. *Open Review of Educational Research*, *4*(1), 118–133. https://doi.org/10.1080/23265507.2017.1354313

Soeharto, S. (2021). Development of A Diagnostic Assessment Test to Evaluate Science Misconceptions in Terms of School Grades: A Rasch Measurement Approach. *Journal of Turkish Science Education*, *18*(3), 351–370. https://doi.org/10.36681/tused.2021.78

Soury-Lavergne, S., & Maschietto, M. (2015). Articulation of spatial and geometrical knowledge in problem solving with technology at primary school. *ZDM*, *47*(3), 435–449. https://doi.org/10.1007/S11858-015-0694-3

Suherman, Prananda, M. R., Proboningrum, D. I., Pratama, E. R., Laksono, P., & Amiruddin. (2020).

Improving Higher Order Thinking Skills (HOTS) with Project Based Learning (PjBL) Model Assisted by Geogebra. *Journal of Physics: Conference Series 1467*, *1467*(1), 012027. https://doi.org/10.1088/1742-6596/1467/1/012027

Sumintono, B., & Widhiarso, W. (2015). *Aplikasi Pemodelan Rasch pada Assessment Pendidikan*. Trim Komunikata.

ten Braak, D., Lenes, R., Purpura, D. J., Schmitt, S. A., & Størksen, I. (2022). Why do early mathematics skills predict later mathematics and reading achievement? The role of executive function. *Journal of Experimental Child Psychology*, *214*, 105306. https://doi.org/10.1016/J.JECP.2021.105306

Utami, N. W., Sayuti, S. A., & Jailani, J. (2021). Indigenous artifacts from remote areas, used to design a lesson plan for preservice math teachers regarding sustainable education. *Heliyon*, *7*(3). https://doi.org/10.1016/j.heliyon.2021.e06417

Uzun, Z., & Öğretmen, T. (2021). Test Equating with the Rasch Model to Compare Pre-test and Post-test Measurements. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, *12*(4), 336–347. https://doi.org/10.21031/epod.957614

Vahey, P., Kim, H.-J., Jackiw, N., Sela, H., & Knudsen, J. (2020). From the static to the dynamic: teachers' varying use of digital technology to support conceptual learning in a curricular activity system. *ZDM* , *52*(7), 1275–1290. https://doi.org/10.1007/S11858-020-01182-6

Velázquez, F. del C., & Méndez, G. M. (2021). Application in augmented reality for learning mathematical functions: A study for the development of spatial intelligence in secondary education students. *Mathematics* , *9*(4), 369. https://doi.org/10.3390/MATH9040369

Wyse, A. E., & Mapuranga, R. (2009). Differential Item Functioning Analysis Using Rasch Item Information Functions. *International Journal of Testing*, *9*(4), 333–357. https://doi.org/10.1080/15305050903352040

Yaqoob, I., Hashem, I. A. T., Ahmed, A., Kazmi, S. M. A., & Hong, C. S. (2019). Internet of things forensics: Recent advances, taxonomy, requirements, and open challenges. *Future Generation Computer Systems*, *92*, 265–275. https://doi.org/10.1016/J.FUTURE.2018.09.058