How the Novelty of Students' Answers in Solving Mathematical Problems?

Retno Widyaningrum[1,2,*], Badrun Kartowagiran[1], Jailani[3]

[1]Department of Educational Research and Evaluation, Yogyakarta State University, Indonesia
[2]Faculty of Tarbiyah and Teacher Training, IAIN Ponorogo, Indonesia
[3]Department of Mathematics Education, Faculty of Math and Science, Yogyakarta State University, Indonesia

Abstract The impact of giving a thinking freedom opportunity is the diversity of students' answers in solving mathematical problems. The smaller the similarity of students' answers compared to all the answers of students in a class, it means that they have higher creative thinking abilities than other students. It is based on a unique answer compared to the others or in other words it has a higher novelty value than the other answers. The objective of this study was to describe the attainment of the novelty of students' answers in solving mathematical problems. The measurement of novelty of students' answers was carried out on 1,002 students from six junior high schools in Indonesia who had different levels of student's input. The scoring system on the response to the completion of multiple solution tasks was by using the four categories of polytomous in which it is based on the Partial Credit Model (PCM). The data were analyzed by using Quest and Parscale program. The result of this study showed that the students have good skills in solving mathematical problems with new ideas on several items. The items referred to some abilities, such as resolving algebra problems involving money operations, designing a flat shape that has specific characteristics, arranging numbers based on certain patterns, and arranging numbers based on certain patterns. The novelty of this research is that this study considers every answer from students or provides a thinking freedom opportunity for students so that mathematical problems are not absolutely solved with one answer.

1. Introduction

A mathematics teacher ideally can develop students' potential, albeit with different characteristics. The results of studies from the [1] explain that teachers can develop students' capacity to connect mathematical ideas and a more in-depth understanding of mathematics through problem-solving with multiple solutions.

Solving mathematical problems with different answers facilitates students to hone maturity in thinking. Students can ensure the truth of connecting mathematical understanding that is used to solve existing problems. It is based on different solutions that can facilitate the connection of the problem at hand to the element of knowledge students have, thereby strengthening the network of related ideas [2].

Several existing studies have been carried out by [3]; [4]; [5]; [6]; [6]; [7]; [8] emphasizing on the constructs and instruments of creativity both in mathematics and other fields. From those studies, it is interesting to follow up concerning schools in Indonesia that have a diversity of students' initial abilities. Schools with complete facilities and good national exam scores will be excellent choices for parents to send their children to school. The creativity of

students who have been accepted at Indonesian schools has never been studied. This intends to explore whether the level of student's creativity is in line with the school level.

One of the subjects taken into account for the reference to graduating students who are studying in Indonesia is mathematics. This shows that mathematics is a science that is considered essential to be studied by students in Indonesia. Therefore, in this study, creativity refers to student's creativity in solving mathematical problems.

[9] propose two main criteria for mathematical creativity: the creation of new knowledge and the ability to solve problems with various correct answers. New knowledge, in this case, can be related to the modification of knowledge construction that has generally been learned by the teacher in the classroom or knowledge that did not yet exist while the ability to solve problems with various things can be a variety of correct answers that are solved using different ways or different results. Different results can be solved in the same way or in different ways. Whereas [10] connects mathematics creativity with students' ability to solve problems with unstructured or diverse answers.

Students' ability to solve problems through different ways or results can be stimulated using problems that can be solved by multiple-solution tasks. A multiple-solution task is a task in which a student is explicitly asked to solve a mathematical problem differently [11]. Whereas [7] defines multiple-solution tasks is not just a way but can also be a different result. The use of test items to measure creativity through the diversity of results is also carried out by [12] to measure the dominant factors forming a person's creativity. If a test taker's answer has less similarity compared to the other test takers' answers, it means that he/she has better creativity. This is in line with the opinion of [13] that one way to measure a person's creativity is through the degree of similarity in the answers or novelty of the answers produced by someone to solve mathematical problems. The tests used to measure student's creativity are limited to the diversity of results different or novelty of students' answers in solving mathematical problems. Based on the description above, this study aims to reveal the novelty of the answers generated by students in junior high schools in solving mathematical problems.

In Indonesia, the measurement of students' creativity in mathematics has not been carried out explicitly [28]. The measurement of mathematical creativity in Indonesia is integrated with exercises in students' textbooks. This is contradictory with the importance of creativity as one of the skills to face the 21st century. Thus, this study was conducted with the aim of revealing students' creativity in solving mathematics problems of Indonesian students.

2. Materials and Methods

2.1. Type of Research

This study consisted of two stages, namely 1) test

planning, and 2) measurement. The preparation of the test in this study is called the test construction method because it constructs a test to measure a particular latent through the underlying dimensions [14]. The construction of the test required a scheme to determine the content to be measured if the test was the norm-referenced test (designed to differentiate abilities among students). The test produced at the planning stage is intended to get an instrument to measure novelty of students' answers in solving mathematical problems.

2.2. Research Stages

The stages of this study are as follows 1) goal setting, 2) determining the novelty variable of students' answers in solving math problems, 3) reviewing students' creativity theory in mathematics, 4) determining the blueprint of test items, 4) item input, 5) scoring rubric, 6) assembling the items into instruments, 7) limited trials, 8) revisions, 9) expert validation, 10) conducting tests on research subjects, 11) data analysis, 12) reports.

2.3. Sample and Data Collection

The population in this study was all junior high school students throughout Ponorogo Regency, East Java, Indonesia. The determination of the Junior High School category is based on the mathematics mean score of national exams in 2019 for all Public and Private Junior High Schools. The school categorization reference in this study is based on the results of the 2019 national exam score of mathematics subject. Schools were categorized as having a high mathematics learning achievement category if their mean score was more than $\bar{X} + \frac{1}{2}s$; schools were categorized as having a medium mathematics learning achievement category if their mean score was more than or equal to $\bar{X} - \frac{1}{2}s$ and less than or equal to $\bar{X} + \frac{1}{2}s$; and schools were categorized as having low mathematics learning achievement category if their mean score was less than $\bar{X} - \frac{1}{2}s$. $\bar{X}$ is the mean score of all respondents, and s is the standard deviation of the mean score of all respondents [16].

Samples in the try-out and measurement activities were taken by stratified cluster random sampling. The sample size in the try-out phase was 506, while the measurement activity was 1002 students. The selected schools were confirmed based on the school's popularity and the location of the school. School's popularity is based on the number of graduates in 2019 that were accepted into senior high with excellent score of national exams. School locations represent locations in the village to the district level. This was done in order to obtain a significant disparity sample.

2.4. Analyzing of Data

Data analysis, at the test preparation stage, is an analysis

of the empirical and theoretical validity (expert review), which is to test that the prepared novelty test includes practical and useful provisions. The data collected is quantitative data obtained from rating scores with a Likert scale filled by expert validators. Qualitative data, at this stage, was also obtained based on descriptive suggestions and comments obtained from experts.

Quantitative data, in the form of validation scores, were analyzed using V-Aiken formula [17], to calculate the index of content validity of novelty test products. The practicality and effectiveness analysis of the novelty test were analyzed by reviewing the mean score of the validator's assessment, while the qualitative data in the form of comments, criticisms, and validator's suggestions were analyzed with descriptive techniques obtained from the same validity sheet.

The try-out data in the form of a score with a Partial Credit Model (PCM) scale were analyzed using the QUEST and Parscale programs. The validity of this study is the construct validity because the test results are expected to reveal a trait or construct of novelty test measured from the construct test induction. Empirical validity based on criterion-related validity (concurrent and predictive) cannot be fulfilled, so that fit item to the model based on Item Response Theory (IRT) was used. IRT has been pioneered by Thurstone since 1925 [18]. The use of IRT was because of the nature of the test in a broad scope so that it is based on the norm-referenced test.

Testing for the determination of the overall test fit for the model as proposed by [19] was by looking at the mean score of the INFIT Mean of Square (mean INFIT MNSQ) along with the standard deviation or by looking at the mean score of the INFIT t (mean INFIT t) with their standard deviation. If the mean score of INFIT MNSQ is close to 1.0 with standard deviation approaching 0.0 or if the magnitude of the mean score of INFIT t (mean INFIT t) is close to 0.0 with standard deviation approaching 1.0, then all test items fit with the model.

Testing the determination of the fit of each item against the model follows the rules set by [19]. An item is fit with the model if the MNSQ INFIT value is in the range of 0.77 to 1.30. The range of INFIT MNSQ values limits the distribution of calibrated scores and is still in the leptokurtic curve, which reflects that it is still in a unity condition. Items that do not fit for the unused model, while those that are less fit (i.e., those with MNSQ INFIT values not too far from the range of 0.77 to 1.30) were reviewed for revision.

3. Result and Discussion

3.1. Result

Test Planning

Based on the test objectives, relevant research reviews,

theoretical studies, learning continuum in this study refers to the results of the publication of the [20] and those that are relevant to Minister of Education concerning the content standards for mathematics subjects studied in Indonesia compiled a blueprint to measure the novelty of students' answers in solving mathematical problems as follows.

Table 1. Blueprint of novelty of students' answers in solving mathematical problems Test

No.	Component	Question Indicator
1	Break Patterns	Arranging numbers based on specific patterns
2	Thinks outside the box	Determining the alternative composition of numbers based on algebraic problems
3	Innovative	Designing various plane figures that have specific characteristics
4	Analysis of the problem	Solving the problem of geometry with certain sized materials
5	Perceptive	Compiling various data that meets certain requirements

Source: adaptation of [21]

The results of the validation were analyzed using *V'aiken* to determine the level of content validity. The determination of the level of content validity was based on *V'Aiken's* price. Data obtained from the expert response patterns resulting from the completion of an item is considered valid if it has a value> 0.75. The price determination of 0.75 is based on the *Aiken* table [22]. The results of the test of content validity based on the price of *V'aiken* obtained 8 valid items. The eight valid items were arranged into 2 test kits. The two test kits consisted of 2 anchor items and 3 different items.

The try-out of the assessment instrument was aimed at checking the validity of the construct and the suitability of the student's novelty measurement model as a latent variable. The try-out results were analyzed with confirmatory factor analysis (CFA) through LISREL 8.51 program for windows. Based on try-out data, it was obtained the following results.

Table 2. Summary of try-out data

Fit model criteria	Device A	Device B	Conclusion
p-value	0,51045	0,23417	*A and B Fit*
RMSEA	0,000	0,033	*A and B Fit*
GFI	0,99	0,99	*A and B Fit*

Based on Table 2, devices A and B that were tested meet the model fit criteria, so it can be concluded that devices A and B fit the model.

The assessment of the significance of parameters or relationships between variables can be seen from the *t-value*. The significance of the parameters to see the relationship between variables can be seen in the regression equation at the output of *Lisrel* with the calculation of the formula based on [23], as follows.
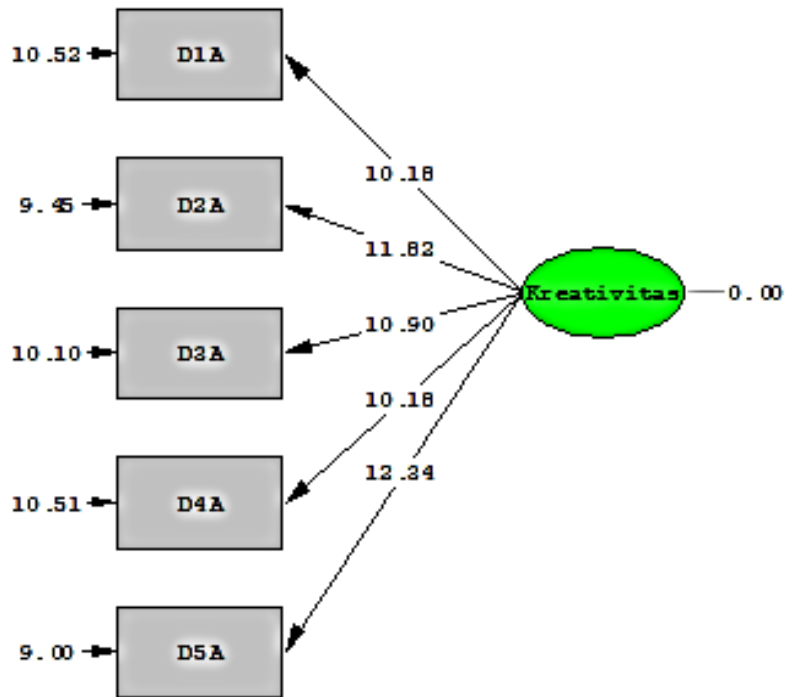
Chi-Square=8.14, df=5, P-value=0.14852, RMSEA=0.045

Figure 1.	t-value of A

Based on figure 1, the value of *t* is higher than *t-table* ± 1.97 at a significance level of 5%, and then the relationship between variables on device A is significant.

Based on figure 2, the value of *t* is higher than *t-table* ± 1.97 at a significance level of 5%, and then the relationship between variables on device B is significant.
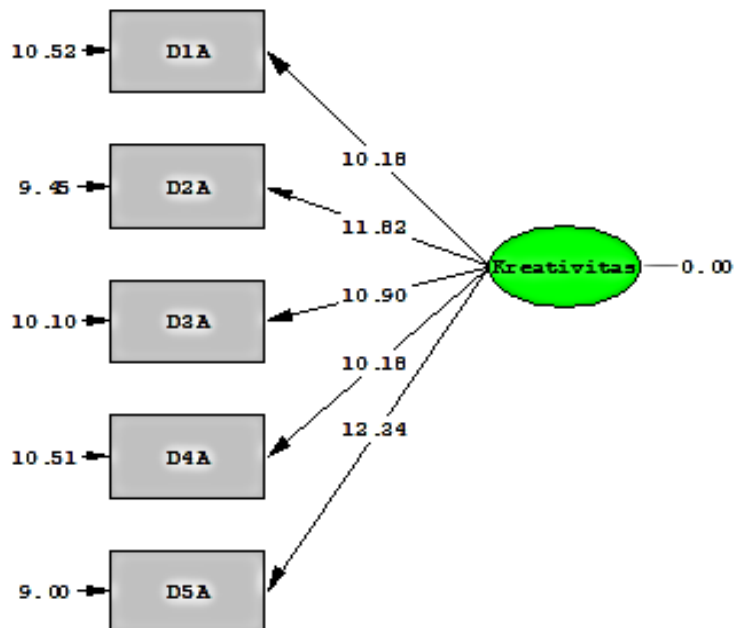
Chi-Square=8.14, df=5, P-value=0.14852, RMSEA=0.045

Figure 2.	t-value of the B test

The case responses in this study were scored based on polytomous scoring with 4 categories, namely categories 1, 2, 3, and 4. The data of the 4 polytomous categories were analyzed using the Partial Credit Model (PCM). Item compatibility analysis with PCM was done through the Quest program. Item compatibility analysis with PCM was seen from the INFIT parameter for mean square (MNSQ) in the output file. The fit value limit that meets the criteria is 0.77 to 1.30 [19].

Based on the results of the analysis of try-out data with the Quest program, the analysis shows that all test items from devices A and B have an MNSQ INFIT value of 0.77 to 1.3, which means that all test items fit with the PCM model. The following figure 3 is a map of the match items from devices A and B using the Quest program obtained as follows.

Figure 3 is the result of the analysis of try-out data for devices A and B, according to INFIT MNSQ PCM. Items 1 and 2 are anchors, items 3 to 5 are part of device A, and items 6 to 8 are part of device B. Based on Figure 3, each item meets the criteria of [19], which is 0.77 <INFIT MNSQ <1.30, so it can be concluded that each item fits with PCM.

A summary of the results of estimated items and tests in the try-out activities is presented in Table 3 below.

Table 3. Summary of Item and Test Result Estimates in Try-out Activities

No.	Description	Estimated Items	Estimated Case
1.	Reliability of estimate	0,87	0,79
2.	The mean score and standard deviation of INFIT MNSQ	0,99 ± 0,07	1,00 ± 0,12
3.	The mean score and standard deviation of OUTFIT MNSQ	1,00 ± 0,14	1,00 ± 0,17
4.	The mean score and standard deviation of INFIT t	-0,10 ± 1,40	-0,02 ± 1,06
5.	The mean score and standard deviation of OUTFIT t	0,01 ± 1,12	0,06 ± 0,84

Based on Table 3, it can be explained that overall, the items arranged in the test kits were tested to fit with the model. This is indicated by the mean value of the MNSQ mean of 0.99 with a standard deviation of 0.08 fulfilling the fit statistic requirements in the Quest program, which is close to the mean score of INFIT MNSQ 1.0 with a standard deviation of 0.0. Likewise, if based on the mean score of INFIT t -0.10 with a standard deviation of 1.9, it approaches the mean score of INFIT t of 0.0 with a standard deviation of 1.0 [19].


```
=========================================================================================
QUEST: The Interactive Test Analysis System
-----------------------------------------------------------------------------------------
Item Fit                                                              14/10/19 11:57
all on all (N = 506 L = 8 Probability Level= .50)
-----------------------------------------------------------------------------------------
INFIT
 MNSQ         .56      .63      .71      .83      1.00     1.20     1.40     1.60     1.80
---------------+---------+---------+---------+---------+---------+---------+---------+---------
   1 item 1                                       .|*              .
   2 item 2                                       *|               .
   3 item 3                                       *|               .
   4 item 4                                    .   |  *            .
   5 item 5                                    .  *|               .
   6 item 6                                    .  *|               .
   7 item 7                                    .   |    *          .
   8 item 8                                    .  * |              .
=========================================================================================
```
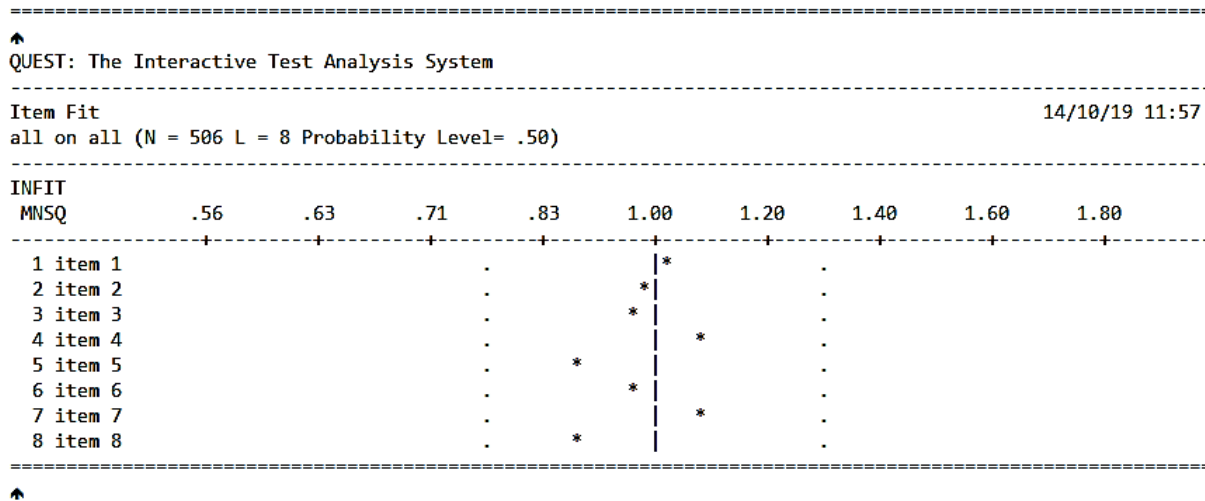
Figure 3. Items on device A and B based on INFIT MNSQ PCM try-out data

The reliability value of the results of the analysis using the Quest program on novelty of students' answers in solving mathematical problems data taken with devices A and B obtained the following results.

```
Summary of item Estimates
=========================

Mean                              .00
SD                               1.06
SD (adjusted)                    1.00
Reliability of estimate           .87
```

Figure 4. The results of the items reliability test on device A and B

```
Summary of case Estimates
=========================

Mean                              .00
SD                                .76
SD (adjusted)                     .62
Reliability of estimate           .79
```

Figure 5. The results of the case reliability test on device A and B

Based on the analysis that has been done, it can be concluded that the device used in the try-out was declared fit with PCM and reliable.

Polytomous scoring has a difficulty level (delta) of more than one. Items with 4 score categories 1, 2, 3, and 4 have deltas from category 1 to category 2, from category 2 to category 3, and from category 3 to category 4. In general, items with n categories have (n-1) delta. Based on the scoring according to PCM on the students' responses to the test results of the test set of 506 respondents who were analyzed using the QUEST program produced an estimate of the level of delta items as presented in Table 4.

Table 4. Difficulty Index and Delta Test Creativity of Students in Try-out

Items	Difficulty	The Level of Item Difficulty		
		Delta-1	Delta-2	Delta-3
1	0,86	-0,62	1,21	0,85
2	0,95	-0,53	0,43	1,31
3	-0,02	-0,07	-0,28	0,37
4	1,08	1,08	-0,45	1,32
5	1,05	-0,91	0,36	1,10
6	-0,02	-0,97	-0,28	1,17
7	1,07	-0,84	-0,45	1,52
8	1,08	-0,91	-0,36	1,10

It is exemplified in Table 4 that to reach category 1 in point 6 (written δ_{61}) of -0.97, which is the lowest delta, while to reach category 3 after reaching category 2 in point 7, δ_{73} of 1.52 is the highest delta.

Scaling, according to PCM, requires that the size of the delta does not have to be met $\delta_{ik} < \delta_{i(k+1)}$. The estimated results in Table 4 also show that not always fulfilled $\delta_{ik} < \delta_{i(k+1)}$. For example point 5, having $\delta_1 = -0,91$, $\delta_2 = 0,36$, and $\delta_3 = 1,10$, means $\delta_1 < \delta_2 < \delta_3$. Item 1, having $\delta_1 = -0.62$, $\delta_2 = 1,21$, and $\delta_3 = 0,85$, means $\delta_1 < \delta_2 > \delta_3$. This can be interpreted that the ability to enter category 3 is easier than the ability to enter category 2. In other words, the number of tests that can enter up to category 3 is greater than the number of tests that enter into category 2. Thus, the characteristics of test items reviewed through the parameter of the difficulty level of delta grains (δ) meet the scaling requirements according to PCM which requires that the estimated value of delta (δ) is not always $\delta_{ik} < \delta_{i(k+1)}$.

3.2. Measurement Phase

The measurement activity involved 1.002 respondents from each of the 2 schools representing high, medium, and low group representatives.

The analysis shows that all 8 test items have INFIT MNSQ between 0.77 and 1.3 so that all test items meet fit with the PCM model. The reliability coefficient is 0.8. A summary of the results of the Quest program analysis is presented in Table 5.

Table 5. Summary of the Results of Measurement Data Analysis

No.	Description	Estimated Items	Estimated Case
1.	Reliability of estimate	0,83	0,80
2.	The mean score and standard deviation of INFIT MNSQ	1,00 ± 0,06	1,01 ± 0,20
3.	The mean score and standard deviation of OUTFIT MNSQ	1,00 ± 0,06	1,00 ± 0,20
4.	The mean score and standard deviation of INFIT t	-0,20 ± 1,32	-0,03 ± 1,16
5.	The mean score and standard deviation of OUTFIT t	0,02 ± 1,21	0,05 ± 0,74

Map parameters are matching test items with PCM from the measurement data analysis results using the help of the following quest program.

```
QUEST: The Interactive Test Analysis System
-------------------------------------------------------------------------------
Item Fit                                                          26/10/19 16:28
all on all (N =1002 L = 8 Probability Level= .50)
-------------------------------------------------------------------------------
INFIT
 MNSQ        .56      .63      .71      .83     1.00     1.20     1.40     1.60     1.80
-------------+--------+--------+--------+--------+--------+--------+--------+--------
   1 item 1                            .       *  |              .
   2 item 2                            .       *  |              .
   3 item 3                            .          |  *           .
   4 item 4                            .          |  *           .
   5 item 5                            .          |  *           .
   6 item 6                            .         *|              .
   7 item 7                            .         *|              .
   8 item 8                            .         *|              .
===============================================================================
```

Figure 6. Items on devices A and B according to INFIT MNSQ PCM measurement data

Figure 6 is the result of the analysis of measurement data for devices A and B, according to INFIT MNSQ PCM. Items 1 and 2 are anchors, items 3 to 5 are part of device A, and items 6 to 8 are part of device B. Based on Figure 6, each item meets the criteria of Adam & Khoo (1996), which is 0.77 <INFIT MNSQ <1.30, so it can be concluded that each item fits with PCM.

The reliability value of the results of the analysis using the Quest program on novelty of students' answers in solving mathematical problems data was taken in the measurement phase as follows.

```
Summary of item Estimates
===========================

Mean                        .00
SD                         1.25
SD (adjusted)              1.23
Reliability of estimate     .83
```

Figure 7. Item reliability based on measurement data

```
Summary of case Estimates
===========================

Mean                        .14
SD                          .07
SD (adjusted)               .80
Reliability of estimate     .80
```

Figure 8. Case reliability based on measurement data

Based on analytical calculations, the reliability value of the items obtained was 0.83, while the value of the reliability of the tests was 0.80. The results of Cronbach's alpha reliability calculation results, according to [24], can be stated that the items used to measure students' creativity are reliable with very high criteria, and the reliability scores of the tests based on these criteria are stated to be reliable with high criteria.

Based on the analysis that has been done, it can be concluded that the device used in the measurement phase is declared fit with PCM and reliable.

The item specifications assembled in the A and B coded test kits that are analyzed using the Quest program have a difficulty test value between -0.06 to 1.08. The average difficulty of novelty test used at the measurement stage was 0.00 ± 0.5 in the medium category. The difficulty distribution of novelty test of full measurement activities is presented in Table 6.

Table 6. Difficulty Index Items for Novelty Tests at the Measurement Stage

No.	Items	Difficulty Index
1.	Resolving algebra problems involving money operations	0,36
2.	Designing a flat shape that has specific characteristics	0,75
3.	Arranging numbers based on certain patterns	-0,06
4.	Designing a room with certain materials	1,06
5.	Compiling data that meet specific requirements	1,01
6.	Arranging numbers based on certain patterns	-0,06
7.	Designing a room with certain materials	1,04
8.	Compiling data that meet specific requirements	1,08

Based on Table 6, it can be seen that items 3 and 6 have the lowest difficulty, while item 8 has the highest difficulty index. In terms of the level of difficulty, item number 8 in the measurement activity is the most difficult item.

Based on Table 6, it can be explained that the difficulty index ranges from -0.06 to 1.08. The difficulty index is classified as good because, according to [25], the difficulty index of a suitable item is in the range of -2 to +2.

Polytomous scoring has a difficulty level (delta) of more than one. Items with 4 score categories 1, 2, 3, and 4 have deltas from category 1 to category 2, from category 2 to category 3, and from category 3 to category 4. In general, items with n categories have (n-1) delta. Based on the scoring according to PCM on student responses, the measurement results of novelty of student answers against 1.002 respondents who were analyzed using the QUEST

program yielded an estimated level of delta difficulty, as presented in Table 7.

Table 7. Difficulty Index and Delta Novelty Tests on Measurement

Items	The Level of Item Difficulty			
	Delta-1	Delta-2	Delta-3	Delta-4
Resolving algebra problems involving money operations	-0,20	0,82	0,18	
Designing a flat shape that has specific characteristics	-0,13	0,93	0,16	
Arranging numbers based on certain patterns	-0,10	0,20	0,85	-0,05
Designing a room with certain materials	-0,27	0,75	0,38	-0,05
Compiling data that meet specific requirements	-0,14	0,88	0,47	-0,11
Arranging numbers based on certain patterns	0,20	0,09	1,10	-0,19
Designing a room with certain materials	0,23	0,58	0,16	-0,15
Compiling data that meet specific requirements	0,37	0,41	0,18	-0,20

Delta of all items ranges from easy to difficult categories with an average of 0.0 and a standard deviation of 0.05. Exemplified that to reach category 1 in item 4 (written δ41) of -0.27, which is the lowest delta, while to reach category 3 after reaching category 2 in item 6, δ63 of 1.10 is the highest delta.

The novelty of student answers can be estimated using the average percentage of student responses answered correctly on each indicator used to measure. Based on Table 32, it is known that students' responses to the testing activities were more dominant in indicators 1 and 2. In other words, a small proportion of respondents were able to reach category 3.

Table 8. Percentage of average responses that answered correctly at the measurement stage

Indicators	Percentage			
	1	2	3	4
Resolving algebra problems involving money operations	20,06	43,41	35,33	1,20
Designing a flat shape that has specific characteristics	19,76	44,61	33,13	2,50
Arranging numbers based on certain patterns	25,65	24,65	47,31	2,40
Designing a room with certain materials	45,61	39,22	12,28	2,89
Compiling data that meet specific requirements	47,80	40,12	10,88	1,20
Arranging numbers based on certain patterns	29,14	23,35	45,21	2,30
Designing a room with certain materials	40,02	36,03	21,76	2,20
Compiling data that meet specific requirements	38,82	35,63	21,76	3,79

Based on Table 8, it can be explained that the items about number patterns are the easiest so that there are more tests in category 4 than the other categories. Whereas designing space with specific materials and compiling data that meets certain conditions is the most difficult item. The dominance of novelty of student answers in the aspects of problem analysis and perspective are in categories 1 and 2 from category 1 to maximum 4. This means that novelty of student answers in the aspects of problem analysis and perspective is not yet satisfactory.

3.3. Discussion

The process of finding problem solving is significant for the development of creative thinking skills [26]. Students who can see a problem from various points of view or describe a problem for several possible solutions to problems are students who have a divergent mindset [27] mentioned that divergent mindset is one indicator of creativity. Stimulating students to see a problem from various points of view so that students can provide alternative solutions will foster new thinking in solving problems which is one of the main aspects of novelty.

The results of the study show that the students were able to complete the items of resolving algebra problems involving money operations, designing a flat shape that has specific characteristics, and arranging numbers based on certain patterns by using new answers. The ability of students to produce new solutions in resolving algebra problems involving money operations is the impact of involving the context of everyday problems. This is relevant with the results of studies [31], [32], [33], [34]. The ability of students to produce new solutions in designing a flat shape that has specific characteristics is the impact of involving students in the learning carried out by the teacher. Student centered learning can facilitate students to develop their potential optimally. This is relevant with the results of studies [35], [36], [37]. The ability of students to produce new solutions in arranging numbers based on certain patterns is the impact of the maturity of the pre-requisite material in a mathematics lesson so that more complex material can be understood. Student centered learning can facilitate students to develop their potential optimally. This is relevant with the results of studies [38], [39], [40].

The results showed that the novelty of student answers aspects of problem analysis and perspective was not satisfactory. The low ability of students to analyze problems is likely because students are not accustomed to solving actual problems [28] explains that mathematical problems that are genuinely problematic and involve significant mathematics have the potential to provide the intellectual context for novelty of student answers development. Problems given to students should not only be solved by ordinary solutions but can stimulate students to analyze problems from various points of view so that

they can provide new solutions that are different from solutions in general. [29] states that the problems given to students to improve the ability to think are not only sufficient to find out the solutions students can produce but problems that can stimulate students' mindset to produce new ideas or "cognitive jumps."

There are several factors that cause the students' low perspective ability. First, students are not accustomed to responding to open questions. The indicator is that only a small proportion of respondents can answer up to category 4. Second, learning that directs novelty of student answers at the research location is only during enrichment phase. Third, mathematics learning at the research location is dominated by the mastery of concepts through mathematical induction. Thus the opportunity to practice developing novelty of student answers is limited. Fourth, the development of novelty of student answers s is influenced by basic mathematical abilities. This is because school input is based on the level of students' achievement in mastering teaching material at the previous educational level. These results are in line with the research conducted by [30].

4. Conclusions

The study shows that the students were able to complete the items of resolving algebra problems involving money operations, designing a flat shape that has specific characteristics, and arranging numbers based on certain patterns by using new answers. The ability of students to produce new solutions in resolving algebra problems involving money operations is the impact of involving the context of everyday problems.. The ability of students to produce new solutions in designing a flat shape that has specific characteristics is the impact of involving students in the learning carried out by the teacher. Student centered learning can facilitate students to develop their potential optimally. The ability of students to produce new solutions in arranging numbers based on certain patterns is the impact of the maturity of the pre-requisite material in a mathematics lesson so that more complex material can be understood. Student centered learning can facilitate students to develop their potential optimally.

REFERENCES

[1] National Council of Teachers of Mathematics, *Principles and standards for school mathematics*. Reston, VA: Author, 2000.

[2] R. Leikin and A. Levav-waynberg, "Solution Spaces of Multiple-Solution Connecting Tasks as a Mirror of the Development of Mathematics Teachers' Knowledge," *Can. J. Sci.*, vol. 8, no. October 2014, pp. 37–41, 2008.

[3] A. Levav-waynberg and R. Leikin, "The role of multiple solution tasks in developing knowledge and creativity in geometry," *J. Math. Behav.*, vol. 31, no. 1, pp. 73–90, 2012.

[4] E. L. Mann and D. Ph, "Mathematical Creativity and School Mathematics: Indicators of Mathematical Creativity in Middle School Students Mathematical Creativity and School Mathematics: Indicators of Mathematical Creativity in Middle School Students," no. January 2005, 2014.

[5] Z. Vally et al., "Examining the effects of creativity training on creative production, creative self-efficacy, and neuro-executive functioning," *Think. Ski. Creat.*, vol. 31, no. October 2018, pp. 70–78, 2019.

[6] K. Yang, Z. Hong, L. Lee, and H. Lin, "Exploring the significant predictors of convergent and divergent scientific creativities," *Think. Ski. Creat.*, vol. 31, no. October 2018, pp. 252–261, 2019.

[7] M. Zainudin, B. Subali, and Jailani, "Construct Validity of Mathematical Creativity Instrument: First-Order and Second-Order Confirmatory Factor Analysis," *Int. J. Instr.*, vol. 12, no. 3, 2019.

[8] X. Huang, J. C. Lee, X. Dong, F. Education, and H. Kong, "Mapping the factors influencing creative teaching in mainland China: An exploratory study," *Think. Ski. Creat.*, vol. 31, no. October 2018, pp. 79–90, 2019.

[9] S. S. Leung, "On the Role of Creative Thinking in Problem Posing," *Analyses,* vol. 29, no. 3, 1997.

[10] M. S. Chiu, "Approaches to the teaching of creative and non-creative mathematical problems," *Int. J. Sci. Math. Educ.*, vol. 7, pp. 55–79, 2009.

[11] R. Leikin and A. Levav-waynberg, "Development of Teachers' Conceptions Through Learning and Teaching: The Meaning and Potential of Multiple- Solution Tasks," *Can. J. Sci.*, vol. 9, no. June 2013, pp. 37–41, 2009.

[12] Yang and Miller, *Karakteristik Responden.* Jakarta: Erlangga, 2008.

[13] R. W. Weisberg, "On the usefulness of 'value' in the definition of creativity," *Creat. Res. J.*, vol. 27, pp. 111–124, 2015.

[14] K. Schweizer and C. DiStefano, "Principles and methods of test construction," *Psychol. Assess.*, vol. 3, 2016.

[15] L. L. Oriondo and Dallo-Antonio, *Evaluating educational outcomes (test, measurement, and evaluation), 5-th*. Quezon City: REX Printing Company, Inc, 1998.

[16] Budiyono, *Metodologi Penelitian Pendidikan* [Methodology of Educational Research]. Surakarta: UNS Press, 2003.

[17] L. R. Aiken, *Psychological testing and assessment*. Allyn & Bacon, 1985.

[18] B. D. Wright and G. N. Masters, *Rating Scale Analysis*. Chicago: Mesa Press, 1982.

[19] R. J. Adams and S.-T. Kho, *Acer Quest Version 2.1.* Camberwell, Victoria: The Australian Council for Educational Research,1996.

[20] Northwest Evaluation Association, *Idaho State Aligned Learning Continuum*. Northwest, 2013.

[21] A. Andiliou and P. K. Murphy, "Examining variations among researchers' and teachers' conceptualizations of creativity: A review and synthesis of contemporary research," *Educ. Res. Rev.*, vol. 5, pp. 201–219, 2010.

[22] L. R. Aiken, "Three Coeficients for Analyzing the Reliability and Validity Ratings," *Educ. Psychol. Meas.*, vol. 45, pp. 131–142, 1985.

[23] I. Ghozali, *Structural Equation Modelling*. Semarang: UNNES Press, 2014.

[24] E. Widoyoko, *Assessment of learning outcome in the schools*. Yogyakarta: Pustaka Pelajar, 2016.

[25] R. K. Hambleton and H. Swaminathan, *Item Response Theory*. Boston, MA: Kluwer Inc., 1985.

[26] M. A. Runco, *Creativity as an educational objective for disadvantaged students (RBDM 9306)*. Storrs, CT: The National Research Center on the Gifted and Talented, University of Connecticut, 1993.

[27] M. Zainudin and E. Istiyono, "Scientific Approach to Promote Response Fluency Viewed from Social," *Eur. J. Educ. Res*., vol. 8, no. 3, pp. 801–808, 2019.

[28] M. Zainudin and B. Subali, "Capability Estimation of Student's Higher Order Thinking in Mathematics by Using Polytomous," vol. 8, no. 3, pp. 895–903, 2020.

[29] V. J. Shute and B. J. Becker, Innovative Assessment for the 21st Century. Florida: Springer Science, 2010.

[30] F. Özdas and V. Batdi, "A Thematic-based Meta Analytic Study Regarding the Effect of Creativity on Academic Success and Learning Retention," vol. 5, no. 3, pp. 53–61, 2017.

[31] L. Ulandari, Z. Amry, and S. Saragih, "Development of Learning Materials Based on Realistic Mathematics Education Approach to Improve Students' Mathematical Problem Solving Ability and Self-Efficacy," Int. Electron. *J. Math. Educ.,* vol. 14, no. 2, pp. 375–383, 2019.

[32] M. Sudia, "Developing skills resolution mathematical primary school students," *Int. J. Educ. Res*., vol. 2, no. 10, pp. 601–614, 2014.

[33] M. S. Amran, F. M. Kutty, and S. Surat, "Creative Problem-Solving (CPS) Skills among University Students," *Creat. Educ.*, vol. 10, pp. 3049–3058, 2019.

[34] O. Ayu, S. Dewi, and E. R. Winarti, "Students â€™ creative thinking ability in solving problems with double loop problem solving model," J. Math. Educ., vol. 8, no. 2, pp. 111–118, 2019.

[35] S. Krishnan, "Student-Centered Learning in a First Year Undergraduate Course," *Int. J. Learn. Teach. Educ. Res.,* vol. 11, no. 2, pp. 88–95, 2015.

[36] I. Zainol, "Student-Centred Learning In Mathematics – Constructivism In The Classroom," *J. Int. Educ. Res.*, vol. 8, no. February, pp. 319–327, 2019.

[37] S. Learning, L. Study, L. Quality, and L. Outcomes, "The Application of Student Centered Learning Through Lesson," *J. Intensive Stud. Lang. Lit. Art, Cult.*, vol. 2, no. 2, pp. 84–91, 2018.

[38] J. Samuelsson and K. Granstr, "Important Prerequisites for Students' Mathematical Achievement," *J. Theory Pract. Educ.*, vol. 3, no. 2, pp. 150–170, 2007.

[39] C. Deeken, I. Neumann, and A. Heinze, "Mathematical Prerequisites for STEM Programs: What do University Instructors Expect from New STEM Undergraduates?," *Int. J. Res. Undergrad. Math. Educ.*, vol. 6, pp. 23–41, 2020.

[40] J. Evensky, D. Kao, Q. Yang, R. Fadele, and R. Fenner, "Addressing prerequisite mathematics needs — a case study in introductory economics," *Int. J. Math. Educ. Sci. Technol*., vol. 5, no. October 2014, pp. 629–639, 2014.