A DIAGNOSIS OF STUDENTS' ERRORS IN ANSWERING THE MATHEMATICS TEST IN SENIOR HIGH SCHOOL

Evana Gina Shantika
Universitas Nahdlatul Wathan Mataram

Edi Istiyono
Universitas Negeri Yogyakarta

Abstract

This study aims to reveal: (1) the characteristics of diagnostic test items used in mathematics subject of the first grade of senior high school (SHS); (2) the degree of errors from several types of students' errors in answering mathematics test; and (3) the dominant factor causing the students to make mistakes in answering mathematics test. This study used a quantitative approach involving the first graders of eight SHSs in Mataram as the population. The sample was collected by a proportionate random sampling technique, consisting of 350 students for preliminary field testing and 450 students for main field testing. The research instrument was a mathematics diagnostic test, questionnaire, and interview guidelines. The results show that: (1) the diagnostic test instrument meets the qualitative and quantitative content validity, proves empirically fit with Partial Credit Model (PCM), has reliability index of 0.92 (high category), and all items in the diagnostic test instrument are categorized in moderate difficulty; (2) misrepresentation becomes dominant errors if compared with misconceptions, the counting errors, and procedural errors. The errors dominant in the topic domain are Inequalities Linear System Two Variables if compared with Rational and Irrational Inequality One Variable, Equations and Inequalities of Absolute Value, and Linear Equations System Three Variables; and (3) most of the students do not understand how to solve the problem of inequality, determining factor quadratic equations, determining the members of the set completion, problems concerning graphs, and problems that require the capability of language interpretation into the mathematics model.

Keywords: *diagnostic errors, mathematics test, Partial Credit Model*

Contact: *Evana Gina Shantika*

evanashantika@gmail.com

Faculty of Teacher Training and Education, Universitas Nahdlatul Wathan Mataram
Jl. Kaktus No. 1-3, Mataram, Nusa Tenggara Barat, Indonesia

Introduction

The measurement and assessment in education are essential, especially towards students. Measurement is an activity which is conducted to describe the characteristics of the object being measured, and it is stated in numbers. The Regulation of the Minister of Education and Culture of Republic of Indonesia No. 23 of 2016 regarding the educational assessment standard, which is used as the foundation in the assessment of students in elementary and middle school, explains that the purpose of learning outcome assessment is to evaluate the process of learning interest, and the remedial of students' learning outcome continuously.

The result of the assessment which has been conducted can become the determinant of the learning outcome, which is based on a proof of measurement result (Mardapi, 2005, p. 75). The assessment of students' learning outcomes is in the form of the test consisting of daily examination, middle semester examination, final semester examination, and national final examination. The government conducts national examination as a medium to equalize the assessment of students' abilities in all areas of Indonesia.

National examination is one kind of formative assessment. The result of students' learning outcome, on a large scale, aims to investigate the achievement of national education based on the national education standard (Mardapi, 2012, p. 223). All assessments which are conducted aim to improve the quality of teachers in teaching and the quality of students in the cognitive, affective, as well as psychomotor aspects. In consequence, the implementation of national examination is expected to be a consideration and indicator in achieving the goals.

Based on the data of the results of national examination in some state senior high schools in Mataram City, the scores of mathematics tests for the last five years are still below the standard. It is presented in Figure 1.
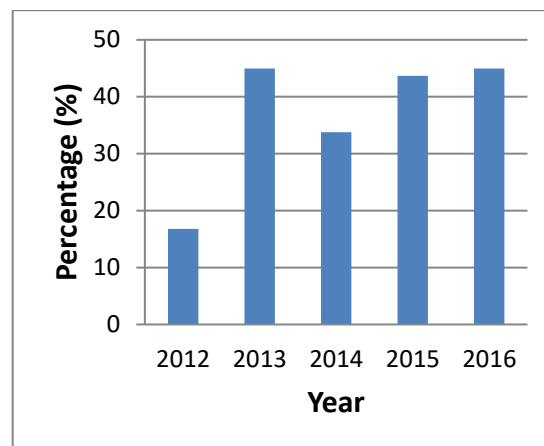

Figure 1. Scores of Mathematics National Examination of State Senior High Schools in Mataram City for the Last Five Years (N < 5.5)

Figure 1 shows that many students obtain the scores far below the national standard, which is 5.5. Based on the analysis of the questions by Isgiyanto (2011), some dominant errors found in studens' answer to the questions, which are categorized into conceptual errors, representation errors, procedural errors, and counting errors.

In realizing the expectation of the national examination, some reformations in the learning process are urgently needed. From the cognitive aspect, students' error in answering the questions were caused by their weakness in some aspects. Therefore, the reformation is expected to improve the result of the mathematics final examination and also can solve the students' weaknesses, especially in certain aspects.

One way to make the learning process more effective is by acknowledging the errors done by students in answering the questions. Acknowledging the errors done by students is one way to analyze the weaknesses in some aspects. Therefore, in the reformation of the learning, teachers should become more direct in emphasizing the students' weaknesses.

The students' errors in answering the questions can be defined as the response of students towards the question item, which is inappropriate with the expected response (Leighton & Gierl, 2007, p. 332). The error which possibly occurs can be identified

using the diagnostic method, especially in analyzing students' strength and weakness in answering questions (Samejima, 1995, p. 402). Diagnosis can be defined as a characteristic or phenomenon in identifying the cause and the decision made through the description or analysis (Embretson, 2007, p. 218). Through a diagnostic method, the information regarding the errors often done by students, the strength and weakness are obtained, and the factors which cause the errors can be acknowledged.

The diagnostic method can be implemented using a polytomous scoring. Polyomous scoring is selected in the diagnosis because it is an item response model that has more than two categories of possibility (Wells, Hambleton, & Purwono, 2008). The difference of polytomous scoring is in the item response where the dichotomous item response has two categories of score, namely the right answer (score 1) and the wrong answer (score 0) (Bond & Fox, 2007, p. 320). It becomes the essential reason why polytomous scoring is used in the diagnostic method.

The more popular model in the initial development of the item response theory is the Partial Credit Model (PCM), the extension of the Rasch model (Retnawati, 2016b, p. 149). In addition, research conducted by Wasis (2011) states that the "partial credit scoring" in physics can produce the estimation of the ability, which is more accurate than the other kinds of scoring, which based on the complexity of every option. It is supported by the research conducted by Isgiyanto (2011) regarding the polytomous scoring with the partial credit model in mathematics learning. The item response theory (IRT) model of PCM polytomous assumes that every question item has the same discrimination index (Embretson & Reise, 2000, p. 125). The use of a diagnostic method based on the polytomous scoring with PCM is expected to give a deeper contribution to diagnosing students' errors.

The consideration of the use of PCM as the extension of the Rasch model, which is a 1-PL model, is because it can utilize the smaller sample than the calibration of the polytomous data using 2-PL or 3-PL model. The characteristic of response in every item follows the PCM, especially in the difficulty level at one stage of the above category, which is not same from each other, so the amount of delta for a stage of the below category and the delta for the stage of the above category is not same to each other (Keeves & Masters, 1999, p. 95). PCM does not require the finishing stage of the test item, which should be ordered or not, and does not need to have the same difficulty level (De Ayala, 1993, p. 175). The test which employs PCM can give information about the difficulty level of students, which can be obtained through the errors they done at every different difficulty level.

Based on the aforementioned description, research to diagnose the students' error in answering mathematics questions is an essential thing, because this research can provide diagnostic information which is needed by teachers in conducting some improvements to obtain the better result for students, especially in mathematics learning. The diagnosis result can also provide information for students to analyze their mathematics skills and motivate them to improve their weak aspects.

Research Method

This study was descriptive-quantitative research based on the polytomous scoring with the Partial Credit Model (PCM) in analyzing the result of the diagnostic test. This research was conducted at eight state senior high schools in Mataram City at the end of the academic year of 2016/2017. The data sampling was conducted in May until June 2017.

The population of this research was 3,598 first graders of those eight senior high schools. The sample was taken through proportionate random sampling. The trial sample comprises of 350 students, and the measurement sample comprises of 450 respondents based on the formula of Krejcie and Morgan (Wagiran, 2013, p. 174).

The results of the diagnostic test were analyzed considering the scoring rubric using polytomous scoring with PCM. Then, it was described to draw a conclusion regarding the characteristic of the mathematic diagnostic test item in the first semester of the first-year students, the percentage of students' error level in various error types determined, and the materials causing errors. The results of questionnaire and interview were described to draw conclusions related to the factors causing students' errors in answering mathematics questions.

The data collection technique used is documentation to obtain data on students' national examination results in mathematics at Mataram City. It is also used to obtain the data of competence and material used in the design of the test. The second technique used is a diagnostic test to diagnose students' errors in answering the questions.

The first step in composing the diagnostic test is determining the aim of the test, that is, to diagnose students' errors in answering math questions in the first year of senior high school. The materials tested are the equation and inequality from the absolute value of the linear form of one variable, rational inequality and irrational one variable, the system of linear equations of three variables, and also the system of inequality of two variables (linear-squared and squared-squared). The materials had been adjusted to the K-13 curriculum used in some state high schools in Mataram.

The next stage is composing the test's outlines, containing the main material, core competence, basic competence, sub material, indicator, types of error, cognitive aspect, and question items. The indicator selected was the one suitable to the domain and ability aspect decided. The diagnostic test item is in the form of multiple choice with every question item having five alternative options with a score of 1-5. Scoring was adjusted to the error level. The distractors were composed in the alternative options through observation in some senior high schools at Mataram. The time given to finish the diagnostic test is 90 minutes.

Based on the outlines, there are 30 indicators in which every indicator is represented by one question. Those questions comprise of 10 question items from the domain of equation and inequality of absolute values in the linear form of one variable, eight items from the domain of rational inequality and irrational variable, four items from the domain of the linear equation system of three variables, and eight items from the domain of two-variable inequality system (linear-square and square-square).

The composing of the test questions was conducted through expert judgment by three experts consisting of one lecturer of mathematics education, one lecturer of educational research and evaluation, and one mathematics teacher of a senior high school in Yogyakarta. The expert judgment was conducted quantitatively and qualitatively. Quantitatively, every test item was scored and analyzed using the Aiken formula. The score of every test item in the validation sheet is between 1 until 4. The expert judgment was also conducted qualitatively in the form of a conclusion of the item with the criteria based on some aspects, such as material, construction, and the language used. The third data collection technique used in this research is a questionnaire and interview to obtain information regarding the factors that cause errors of students in answering the questions, seen from the aspect of competence needed by students in answering the questions.

Findings and Discussion

The Characteristic of Question Item

Validity and Reliability

The content validity is related to the rational analysis of the domain to be measured to find out instructional representation with the ability to be measured (Retnawati, 2016a, p. 158). The first-year semester one mathematical diagnostic test instrument fulfilled the content validity qualitatively with expert judgment and quantitatively obtained the Aiken index of 0.848 in the high catego-
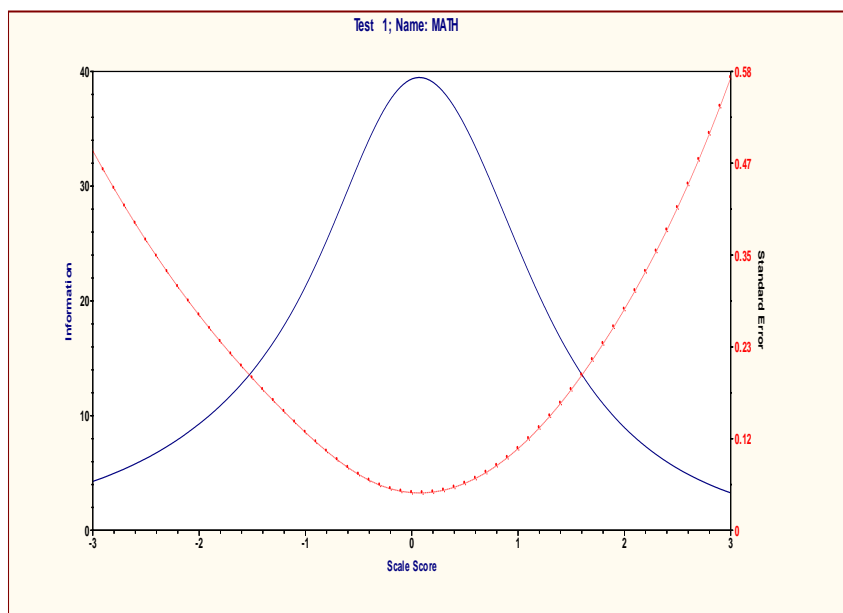
Figure 2. The Information Function and SEM

ry (Retnawati, 2016b, p. 19). The estimated function of the test information and the calculation of the standard error of measurement is shown in Figure 2.

Based on Figure 2, the information and SEM function intersect on a scale of -1.6 and +1.6. It means that the diagnostic test that has been prepared is suitable for test-takers in the ability range of -1.6 to +1.6. The estimated reliability results of 0.92 are classified as a high category.

Item Response Theory

The analysis in this study uses the IRT approach. Before an analysis using the IRT approach, there are several assumptions that must be met, namely unidimensional and local independence (Hambleton, Swaminathan, & Rogers, 1991, pp. 9–12; Mardapi, 2012, p. 201). The unidimensional assumption test is done by exploratory factor analysis. Before further factor analysis (Eigen Value and Scree Plot), the sample adequacy test is performed using Barlett's-Test as presented in Table 1.

Based on Table 1, the result for the KMO MSA diagnostic test instrument >0.5 is 0.895 with a significance value of the Bartlett's Test of Sphericity <0.5, 0.000. It shows that the number of the students' responses have met the requirements for factor analysis. The requirements for the factor analysis are the KMO MSA value >0.5 and the significance value of the Bartlett test <0.05 (Hair, Anderson, Tatham, & Black, 2006, p. 115). In addition, there are also things that need to be considered, namely the MSA value per item in the Anti-Image Correlation matrix. Based on the Anti-Image Correlation matrix, it is known that there is no item that has an MSA value <0.5, so there is no aborted item, and there is no need to re-test (Susetyo, 2015, p. 70).

Table 1. The Result of KMO MSA Test and Bartlett's Test of Sphericity

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		0.895
Bartlett's Test of Sphericity	Approx. Chi-Square	3255.986
	Df	435
	Sig.	.000

Unidimensional proof can be performed by looking at the Eigen value or scree plot that is formed based on the Eigen value. Furr states that a test is said to be unidimensional if components 1 and 2 in the scree plot have a considerable distance, as shown in Figure 3 (Furr & Bacharach, 2008, p. 74).
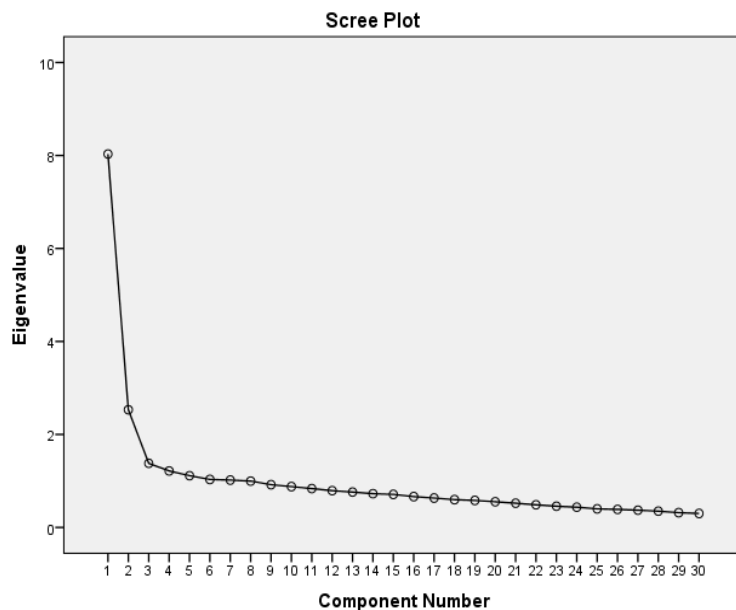
Figure 3. Scree Plot of Eigen Test Value of Mathematics Diagnostic Test

Figure 4. The Output Quest of Infit MNSQ Value in the Determination of Fit Model

Based on Figure 3, the scree plot shows that factor 1 looks steep because it has a great distance with factor 2, while factor 2 to factor 3 and the forth starts to ramps because it has a very close distance. In addition, the scree plot also shows that there is only one steepness, namely in factor 1 to factor 2. It corroborates the results which show that the diagnostic instrument of mathematics for the first year of high school is unidimensional.

The assumption of local independence is automatically proven after being proven by the unidimensionality of the test participant's response data. It means that the test participant's answer to an item has no relationship with other items in a test set (Retnawati, 2014, pp. 3–7).

Model Suitability (Goodness of Fit (GoF))

Determination of overall item fits the model using the Quest program (Adams & Khoo, 1996, p. 90), which is based on the magnitude of the average value of INFIT Mean of Square (INFIT MNSQ) and standard deviation or the average value of INFIT Mean of INFIT t. The Quest program output for fit model analysis is presented in Figure 4.

Based on Figure 4, it can be seen that for each MNSQ INFIT value, it is in the range of values of 0.77 to 1.30. Thus, it indicates that each item on the mathematical diagnostic test instrument for the first graders of senior high school fits the model.

The Level of Test Item Difficulty

Theoretically, the difficulty level index (b_i) lies between -2 and +2. Difficulty value approaches -2 identifies that the item is easier. Moreover, if the difficulty value approaches +2, it shows that the item is a more difficult item (Mardapi, 2012, p. 205). The results of the analysis of the item's level of difficulty are summarized in Table 2.

Table 2. The Summary of Analysis Result of Item Difficulty Level

Question Category	Item Number	Total	Exp.
Concept	1, 4, 7, 11, 14, 19, 23, 26	8	Medium
Procedure	2, 5, 8, 12, 15, 20, 24, 27	8	Medium
Counting	3, 6, 13, 16, 21, 25, 28	7	Medium
Representation	9, 10, 17, 18, 22, 29, 30	7	Medium
Total		**30**	

Table 2 shows that the difficulty level of each item on the diagnostic test instrument is in the moderate category. A good point of difficulty is located at the -2 interval b≤2 interval (Hambleton et al., 1991, p. 13).

Error Level in Every Error Type

Error Concept

Students who make a conceptual mistake means the students are wrong in answering questions about the concept categories that are included in scores 1, 2, 3, and 4. Items included in the concept category are items 1, 4, 7, 11, 14, 19, 23, and 26, which in answering, they only need an understanding of the concept. Conceptual understanding is the ability to use new ideas depending on how to connect to previous ideas and processes. In learning mathematics, students can connect between new material and material that was previously received (prerequisite) (Hasselbring, Lott, & Zydney, 2005). The distribution of the students' scores in answering the question of concept categories is presented in Figure 5.
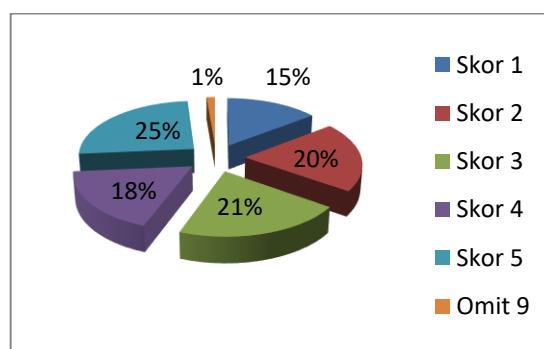

Figure 5. The Distribution of Score in the Conceptual Question

Based on Figure 5, it can be seen that students who made a conceptual error that is getting a score of 1, 2, 3, and 4, are 74%. Students who were correct in answering questions about the concept category or those included in the score of 5 are 25%, while 1% did not answer the questions.

Procedural Error

Students who make mistakes in answering items number 2, 5, 8, 12, 15, 20, 24, and 27 mean that they have made a procedural error, where these items are included in the procedure category, which in answering them, it requires proper understanding of procedures. These students are

wrong in answering the question of procedure categories in scores 1, 2, 3, and 4. This procedural error occurs because students do not know the procedures required to carry out operations accurately, even though students are able to identify the correct operation or sequence of operations (White, 2005, p. 15). The form of questions in the procedure category is to choose the right procedure or not the right one in solving a problem. The distribution of students' scores in answering the question of procedure categories is presented in Figure 6.
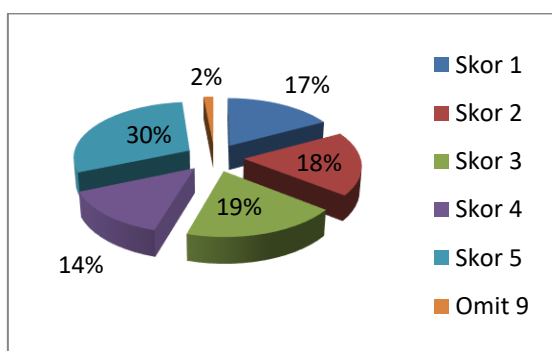

Figure 6. The Distribution of Score in Procedural Question

Figure 6 shows that students who made a procedural error that is included in scores 1, 2, 3, and 4 are 68%. Students who are correct in answering questions about the procedure category or who belong to a score of 5 are 30%, while 2% do not answer the questions.

Counting Error

Items included in the counting category are items number 3, 6, 13, 16, 21, 25, and 28 where the answer choice distractor in the calculation problem category is arranged by placing an error on the calculation error in the arithmetic operations of addition, subtraction, multiplication, and division. Indicators of difficulty in numeracy skills are difficulty in computing, difficulty in manipulating operations, and not rechecking the results of calculations (Retnawati, Pardi, & Prastowo, 1994, pp. 12–13). Students who make a numeric mis-

take means the students are wrong in answering the numeracy category questions in scores 1, 2, 3, and 4. The distribution of students' scores in answering numeracy category questions is presented in Figure 7.
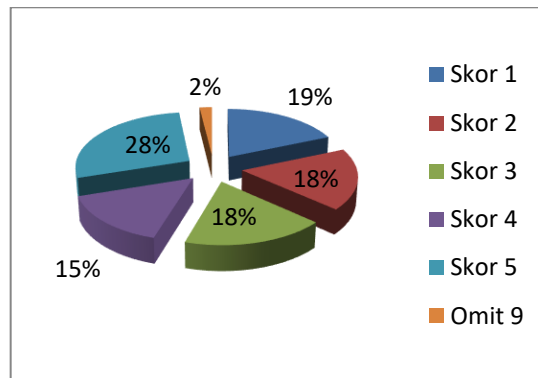

Figure 7. The Distribution of Score in Counting Questions

Based on Figure 7, students who make mistakes in counting are those included in scores 1, 2, 3, and 4 by 70%. Students who were correct in answering questions on the numeracy category or who included in the score of 5 were 28%, while 2% did not answer the questions.

Representation Error

The form of category representation questions is about a story or application problem. Indicators of weakness in language interpretation are difficulty in expressing everyday language in mathematical language, difficulty in interpreting graphs or tables into mathematical language, and difficulty in expressing mathematical language in everyday language (Retnawati et al., 1994, pp. 12–13). The items included in the category of representation are items 9, 10, 17, 18, 22, 29, and 30, which in answering, it requires the ability to represent the problem language into mathematical language. Students who make a misrepresentation means that students are wrong in answering questions about the categories of representtation, including in scores 1, 2, 3, and 4. The distribution of students' scores in answering questions about representation is presented in Figure 8.
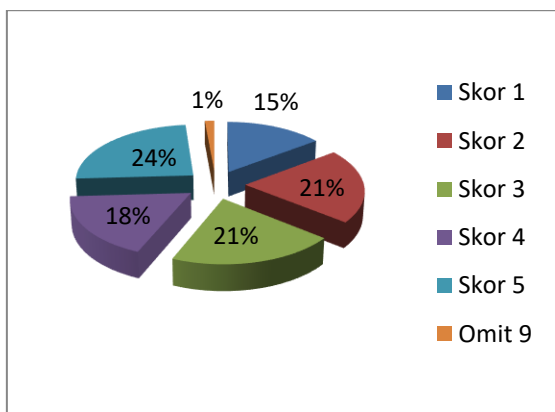
Figure 8. The Distribution of Score in
Representation Questions

Figure 8 shows that students who made a misrepresentation were in scores 1, 2, 3, and 4 by 75%. Students who were correct in answering questions about the category of representation or included in the score of 5 were 24%, while 1% did not answer the questions. The comparison of error rates for each type of error can be seen in Figure 9.
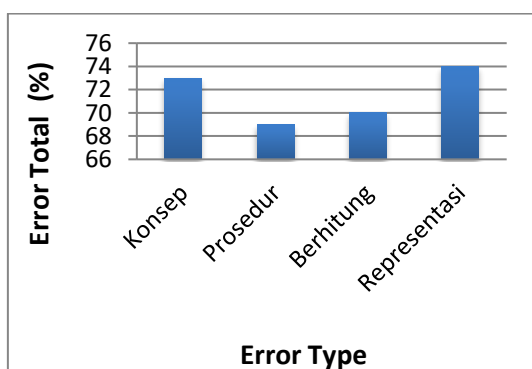

Figure 9. Error Level in Every Error Type

Based on Figure 9 and exposure to the level of errors in each type of error, it can be seen that the errors of students in answering math problems for the first year of high school in the first semester dominantly lie in the error of representation with the highest percentage of errors of 75% among the other percentage of errors. The order of error rates of the types of errors starting from the highest are the error of representation, then followed by concept errors, counting errors, and the smallest percentage is in procedural errors.

The Error Level in Every Material Tested

The Equation and Inequality of Absolute Value of One Linear Variable Form

The item questions which are included in the Equality and Inequality Domains Absolute Value Form of Linear One Variable are items number 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10. The distribution of students' scores in answering the questions about the Equation domain and the Inequality of Absolute Value of Linear Forms of One Variable is presented in Figure 10.
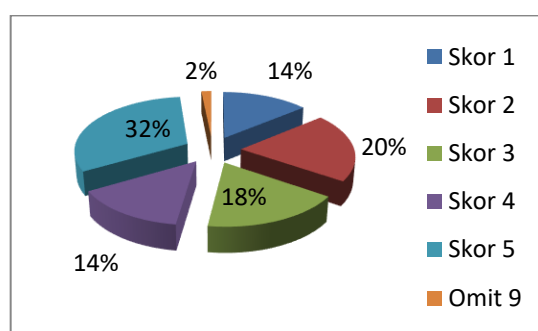

Figure 10. The Distribution of Score in the Questions of Equation and Absolute Inequality Value Domain

Based on Figure 10, it can be seen that students who make mistakes in answering questions about the Equation and Inequality Absolute Value Form of Linear One Variable, namely in scores 1, 2 3, and 4, are 66%. Students who are correct in answering questions about the domain of Equality and Inequality Absolute Value in the Form of Linear One Variable or included in the score of 5 are 32%. Meanwhile, 2% of them do not answer the questions.

The Rational and Irrational Inequality of One Variable

The question items which are included in the Rational and Irrational Inequality domain of One Variable are items number 11, 12, 13, 14, 15, 16, 17, and 18. The distribution of students' scores in answering questions about the Rational and Irrational Inequality Variable One variable is presented in Figure 11.
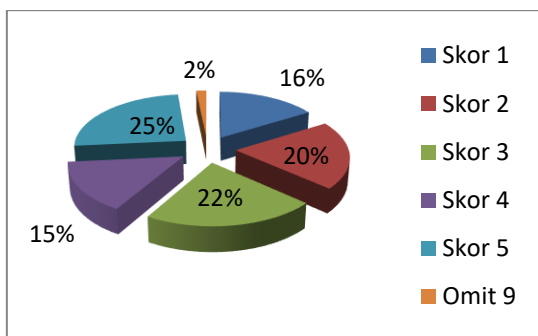
Figure 11. The Distribution of Score in the Questions of Rational and Irrational Inequality of One Variable

Figure 11 shows that students who made mistakes in answering questions about the domain of Rational and Irrational Inequality of One Variable, namely in scores 1, 2, 3, and 4 are 73%. Students who are correct in answering questions about the Rational and Irrational Inequality Variable Variables or those that score 5 are 25%, while 2% do not answer the questions.

Linear Equation System of Three Variable

The items included in the Three Variable Linear Equation System domain are items number 19, 20, 21, and 22. The distribution of students' scores in answering the problem domain of the Three Variable Linear Equation System is presented in Figure 12.
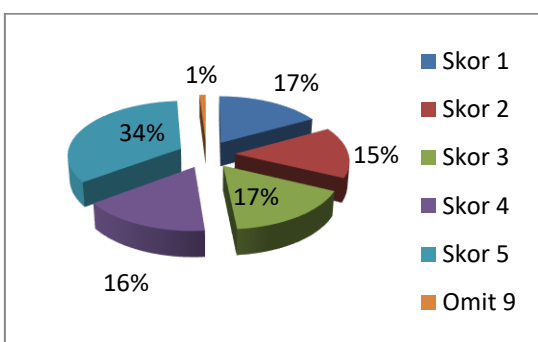

Figure 12. The Distribution of Score in the Questions of Linear Equation System of Three Variables

Based on Figure 12, students who make mistakes in answering the problem domain of the Three Variable Linear Equation System are in scores 1, 2, 3, and 4

by 65%. Students who are correct in answering questions about the domain of the Three Variable Linear Equation System or those included in the score of 5 are 34%, while 1% does not answer the questions. Research conducted by Blanco and Garrote (2007, p. 228) shows that students who are not able to master the basic concepts in arithmetic will result in these students not being able to solve linear equations.

The Inequality System of Two Variables (Linear-Square andSquare-Square)

The items included in the Two-Variable Inequality System domain (Linear-Squares and Squares-Squares) are items number 23, 24, 25, 26, 27, 28, 29, and 30. The distribution of students' scores in answering the Inequality System domain questions Two Variables (Linear-Squared and Squares-Squares) is presented in Figure 13.
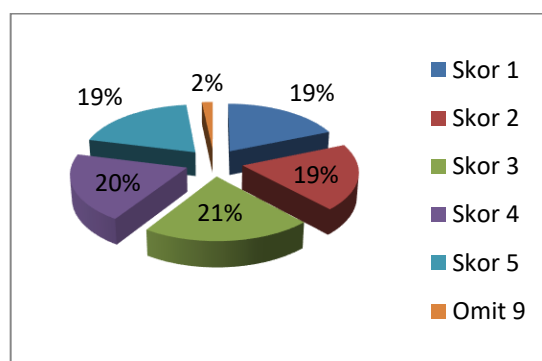
Figure 13. The Distribution of Score in the Questions of Inequality System of Two Variables

Figure 13 presents that students who made mistakes in answering the problem of the Inequality System of Two Variables (Linear-Square and Square-Squared), namely, in a score of 1, 2, 3, and 4 are 79%. Students who are correct in answering domain Inequality Systems Two Variables (Linear-Squares and Squares-Squares) or those included in the score of 5 are 19%, while 2% do not answer the questions. The comparison of error rates for each material is shown in Figure 14.

Based on Figure 14 and exposure to the error rate on each material, the errors of

students in answering math problems for the first graders of senior high school in the first semester dominantly lie in the material of Linear Inequality System Two Variables with the highest percentage of errors of 79% among the highest percentage of errors. The order of the error rate of the material being tested starting from the highest is the error in the Linear Inequality Two Variable System domain, then followed by an error in the Rational and Irrational Inequality domain of One Variable, the Equation domain and the Inequality of the Absolute Form of a Linear Variable, and the smallest percentage of errors is in the domain of the Three Variable Linear Equation System.
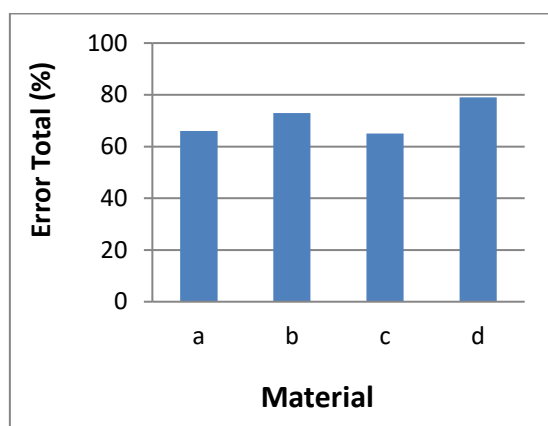

Figure 14. Error Level in Every Material

The Factors Causing Errors when Students Answer the Mathematics Questions

Interview

Through a structured interview conducted on eight mathematics-subject teachers in each of the state high schools in the city of Mataram, some information about the factors causing students to make mistakes in answering math problems is obtained. In terms of the material aspect, students still have difficulty in understanding the linear inequality two-variable system material, which contains material about linear-squared inequality systems and quadratic-inequality inequality systems. This material includes a material with a high degree of difficulty because understanding this mate-

rial requires a deep understanding of quadratic equations and inequalities, both in its operation, determining factors, and graphs about quadratic equations and inequalities. It causes the linear inequality two-variable system is more difficult to understand. Moreover, in solving problems related to the system of linear inequality two variables, it requires a fairly long solution.

The results of research conducted by Kusaeri (2012) shows that the basic operations or basic concepts in arithmetic contribute to the mastery of students in inequality. Understanding the basic concepts of algebra and equality will influence students' mastery of concepts on the concept of inequality.

Previewed from the aspects of the prerequisite material, students are still weak in terms of working on inequality questions. There are still many students who do not understand the problems related to graphics, do not understand the material concepts related to the problems and concepts of prerequisite materials needed, and the students' weak ability in interpreting the intentions of the problem, especially in the case of a story or application.

From the aspect of implementing mathematics learning in the classroom, the teacher explained that students were not too tense and more active in learning mathematics. Many learning methods have been conducted, including learning using lecture and group discussion methods, sometimes giving students worksheets, as much as possible creating a relaxed learning atmosphere by joking occasionally so that students are not too tense in learning mathematics, increasing the number of questions and answers to students during learning so that learning activities are more active, and carrying out a remedial for students who did not pass the test.

Viewed from the aspect of the availability of materials/tools and media for learning mathematics, the teacher explained the availability was quite good, such as books for each student, and LCD installed in each class because they often use power

point media and videos in learning mathematics in class. Power point media are often used in learning mathematics, whereas media for teaching aids in learning mathematics depends on the material being taught whether or not it needs media.

The factors explained by the teachers who are considered to be able to inhibit the mathematics learning in the classroom include (1) students' lack of focus in learning, especially when there are math courses at the seventh and eighth sessions, (2) a large number of students in one class which makes it difficult for teachers to control students, (3) school activities that lead to reduced effective learning time of mathematics so that often, when approaching to the semester exam, many materials have not been completed yet, and at the end of the learning, the classroom is accelerated to be able to finish all the material before the semester ends, (4) there are still many students who are afraid to ask questions, (5) students' interest in learning math tends to be very lacking, and (6) the students' mindset that mathematics is difficult.

The teacher further explained the efforts made to overcome obstacles in learning mathematics in the classroom, including by (1) providing motivation to students to be more enthusiastic in learning mathematics, (2) forming learning groups by appointing one student as a group leader in each group responsible to the group members in order to understand the learning material, (3) giving punishment to students who make learning conditions uncomfortable, (4) providing additional lessons outside of school hours or sometimes asking for additional hours from other subject teachers whose subject material is almost complete, (5) occasionally joking with students to create the fun atmosphere of learning if students begin to look tired and lose the focus, (6) approaching the students outside of the class hours so that when learning in class they are not afraid to express opinions or ask questions directly to the teacher, and (7) changing the mindset about mathematics as a difficult subject.

In addition, if viewed from the aspect of the relationship between teacher and students, the teacher explained that the relationship is considered to be quite good. It is because the teacher realizes that fostering a good relationship between the teacher and students can facilitate learning activities in the classroom. One of the advantages is that when learning, students are not afraid to ask questions directly to the teacher.

Open-Ended Questionnaire

The conclusion obtained from the open-ended questionnaire is about the factors that can cause students to make mistakes in answering mathematics questions, in terms of the ability needed by students to answer the questions correctly. In the material of Equality and Inequality Absolute Value Form Linear One Variable, most students still do not understand the steps of solving absolute inequality values, determining the members of the settlement set, using inequality marks, the steps in drawing graphs of absolute value, and interpreting the language of the problem into mathematical models. There are still many students in the domain of Rational and Irrational Inequality One Variable who do not understand the use of inequality signs, how to determine the members of the settlement set, and how to interpret the language of questions into rational/irrational inequalities. Viewed from the material of the Three-Variable Linear Equation System, most students do not understand how to use the elimination method and interpret the problem language into a mathematical model.

The last material domain is the Inequality System of Two Variables (Linear-Square and Square-Squared), where, in this material, there are still many students who do not understand how to determine the quadratic equation factor, determine linear and quadratic inequality, and determine the intersection of quadratic charts. Mastery of the material in the first semester among the first graders of high school is very closely related to the mastery of the material at the

previous level of education. It is in line with Warren (2003) that students experience the process of thinking transition in junior high school that is the transfer of knowledge needed to solve arithmetic equations (operations involving numbers or numbers) to the knowledge needed to solve algebraic equations (operations involving variables) which bring changes in the shape of the object of study from numbers to variables, equations, and so on.

Conclusion

Based on the results of the study, several conclusions are drawn. First, concerning the characteristics of diagnostic test items, the mathematical diagnostic test instrument for the senior high school first graders in the first semester fulfills the content validity qualitatively and quantatively, and it has been proven empirically fit with the Partial Credit Model (PCM) based on five categories of polytomous data. The instrument reliability estimate is 0.92 in the high category. Based on information and SEM functions, this diagnostic test instrument is very appropriate to be used for test participants in the ability range of -1.6 to +1.6. All items contained in the diagnostic test instrument used in the study have a moderate level of difficulty.

Second, it is concluded that the mistake of representation is the most dominant error made by students in answering math problems compared to concept errors, calculation errors, and also procedural errors. Viewed from the material aspect, it is concluded that the most dominant mistakes made by students are in the domain problems of the Linear Inequalities of Two Variables (Linear-Square and Square-Squared) when compared to the domain aspects of the Rational and Irrational Inequality Problems of One Variable, Inequality Value Absolute, and Three Variable Linear Equation System.

Third, the dominant factor causing students' errors in answering the math problems is in the aspect of the ability needed by students in answering questions. This aspect is elaborated as follows: most students still do not understand how to solve problems about inequality, determine the quadratic equation factor, determine the set members solving, solve problems with graphics, and solve problems that require the ability to interpret the language of problems into mathematical models.

References

Adams, R. J., & Khoo, S.-T. (1996). *Acer quest version 2.1*. Camberwell, Victoria: Australian Council for Educational Research.

Blanco, L. J., & Garrote, M. (2007). Difficulties in learning inequalities in students of the first year of pre-university education in Spain. *EURASIA Journal of Mathematics, Science and Technology Education*, *3*(3), 221–229. https://doi.org/10.12973/ejmste/75401

Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

De Ayala, R. J. (1993). Methods, plainly speaking: An introduction to polytomous item response theory models. *Measurement and Evaluation in Counseling and Development*, *25*(4), 172–189.

Embretson, S. E. (2007). Construct validity: A universal validity system or just another test evaluation procedure? *Educational Researcher*, *36*(8), 449–455. https://doi.org/10.3102/0013189X07311600

Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologist*. London: Lawrence Erlbaum Associates.

Furr, R. M., & Bacharach, V. R. (2008). *Psychometrics: An introduction*. Thousand Oaks, CA: Sage Publications.

Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (2006). *Multivariate

data analysis (6th ed.). Upper Saddle River, NJ: Prentice-Hall.

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. Newbury Park, CA: Sage Publications.

Hasselbring, T. S., Lott, A. C., & Zydney, J. M. (2005). Technology-supported math instruction for students with disabilities: Two decades of research and development. Retrieved June 10, 2017, from LD Online website: http://www.ldonline.org/article/6291/

Isgiyanto, A. (2011). Diagnosis kesalahan siswa berbasis penskoran politomus model partial credit pada matematika. *Jurnal Penelitian Dan Evaluasi Pendidikan*, *15*(2), 308–325. Retrieved from https://journal.uny.ac.id/index.php/jpep/article/view/1099/1151

Keeves, J. P., & Masters, G. N. (1999). Partial credit model. In G. N. Masters & J. P. Kevees (Eds.), *Advances in measurement in educational research and assesment*. Amsterdam: Pergamon-Elsevier Science.

Kusaeri, K. (2012). *Pengembangan tes diagnostik dengan menggunakan model DINA untuk mendapatkan informasi salah konsepsi dalam aljabar*. Doctoral dissertation, Universitas negeri Yogyakarta, Yogyakarta.

Leighton, J. P., & Gierl, M. J. (2007). Defining and evaluating models of cognition used in educational measurement to make inferences about examinees' thinking processes. *Educational Measurement: Issues and Practice*, *26*(2), 3–16. https://doi.org/10.1111/j.1745-3992.2007.00090.x

Mardapi, D. (2005). Pengembangan sistem penilaian berbasis kompetensi. In Himpunan Evaluasi Pendidikan Indonesia (Ed.), *Rekayasa sistem penilaian dalam rangka meningkatkan kualitas pendidikan* (pp. 71–85).

Yogyakarta: Himpunan Evaluasi Pendidikan Indonesia.

Mardapi, D. (2012). *Pengukuran, penilaian, dan evaluasi pendidikan*. Yogyakarta: Nuha Medika.

Regulation of the Minister of Education and Culture of Republic of Indonesia No. 23 of 2016 on Educational Assessment Standard, (2016).

Retnawati, H. (2014). *Teori respons butir dan penerapannya: Untuk peneliti, praktisi pengukuran dan pengujian, mahasiswa pascasarjana*. Yogyakarta: Nuha Medika.

Retnawati, H. (2016a). Proving content validity of self-regulated learning scale (The comparison of Aiken index and expanded Gregory index). *REiD (Research and Evaluation in Education)*, *2*(2), 155–164. https://doi.org/10.21831/reid.v2i2.11029

Retnawati, H. (2016b). *Validitas reliabilitas & karakteristik butir (Panduan untuk peneliti, mahasiswa, dan psikometrian)*. Yogyakarta: Nuha Medika.

Retnawati, H., Pardi, P., & Prastowo, R. (1994). *Identifikasi kesulitan calon guru matematika untuk menyelesaikan persoalan matematika DIY*. Yogyakarta: Universitas Negeri Yogyakarta.

Samejima, F. (1995). A cognitive method using latent trait models: Competency space approach and its relationship with Dibello and Stout's unified cognitivepsychometric diagnosis model. In P. D. Nichols, S. F. Chipman, & R. L. Brennan (Eds.), *Cognitively diagnostic assessment* (pp. 391–410). Mahwah, NJ: Lawrence Erlbaum Associates.

Susetyo, B. (2015). *Prosedur penyusunan dan teknik analisis tes*. Bandung: Refika Aditama.

Wagiran, W. (2013). *Metodologi penelitian pendidikan: Teori dan imlementasi*. Yogyakarta: Budi Utama.

Warren, E. (2003). The role of arithmetic structure in the transition from arithmetic to algebra. *Mathematics Education Research Journal, 15*(2), 122–137. https://doi.org/10.1007/BF032 17374

Wasis, W. (2011). *Penskoran model partial credit pada item multiple true false bidang fisika.* Doctoral dissertation, Universitas Negeri Yogyakarta, Yogyakarta.

Wells, C. S., Hambleton, R. K., & Purwono, U. (2008). Polytomous response IRT models and applications. *Pelatihan Asesmen Pendidikan Dan Psikologi (Psikometri)*, pp. 2–10. Yogyakarta: Program Pascasarjana Universitas Negeri Yogyakarta.

White, A. L. (2005). Active mathematics in classrooms: Finding out why children make mistakes - and then doing something to help them. *Square One, 4*, 15–19.