The final semester test of mathematics subject in vocational high school, how difficult is it?

D A Krisma and H Retnawati
Mathematics Education, Graduate Program, Universitas Negeri Yogyakarta, Indonesia
E-mail: ditaaldila.2018@student.uny.ac.id

Abstract. Progression, result, and completeness of student learning can be assessed through a final semester assessment. In order to make a good assessment, the quality of test items is needed. This study aimed to describe the characteristics of the final semester test items used in “Indonesia” Yogyakarta Vocational High School. This study was a quantitative research with explorative descriptive type. The data obtained from the set of mathematics final semester test items used in grade XI “Indonesia” Yogyakarta Vocational High School along with student response. Documentation was applied for data collection. The set of test items was validated by 3 experts and its characteristics were analyzed with classical test theory. The results showed that the content validation is 0.82 invalid categories. Coefficient reliability is 0.77 at a reliable degree. Test items that have easy difficulty level are 4 items, the medium difficulty level are 21 items, and the difficulty level is difficult as many as 4 items. Item discrimination on item test showed that 21 items are included in the good category and 8 items are not good. The distractor of multiple-choice items worked effectively.

1. Introduction
In the education sector, assessment is important to analyze a measure of successful learning which can measure students' ability in mastering the competencies that have been determined. This is in line with the statement Biggs [1] and Hadi & Elvira [2] that assessment of learning outcomes is one of the efforts made to measure the quality of education with its essential function of measuring the success of learning carried out by the teacher while measuring the success of students in mastering competence. Furthermore, Jabbarifar [3] states assessment is a full range of procedures used to gain information about the progress of learning participated by students through observation, students' performances doing task, and test, moreover to obtain information value judgment formation concerning in learning progress. The results of assessments, educators, educational units, and the government can find out the results of the learning process and can be a benchmark for improving the learning process. The assessment consists of daily assessments carried out by the educators, final semester assessments carried out by the educational units and national examination carried out by the government.

There are several things that must be considered in the assessment. Assessment is a process that includes four basic components, namely measuring daily progress, motivating students to learn, evaluating teaching methods, and knowing students' abilities in relation to evaluating all study groups [4]. The assessment can help to find out something [3], including the extent to which students achieve the learning objectives, what kind of difficulties are faced by students, which students should be referred to counseling, special classes, or improvement programs, and which students have low self-
understanding. Furthermore, the assessment will be more effective if it follows the following principle [3]:

a. Determine clearly what will be assessed as having priority in the assessment process.
b. The assessment procedure must be chosen because of its relevance to the characteristics or performance to be measured.
c. Comprehensive assessment requires a variety of procedures.
d. The use of appropriate assessment procedures requires awareness of its limitations.
e. Assessment is a means to an end, not an end in itself

One of the assessment objectives is the assessment for learning or summative tests. One of summative test used to measure students' abilities is the final semester test in the even semester. This test is an activity carried out to measure student achievement at the end of the even semester. This final semester test is one form of educational assessment at the level of primary and secondary education organized by the education unit. The scope of assessment includes all indicators that represent all basic competencies in the even semester. The final semester test can determine the position of students' ability to master the material. The function of this test also a diagnostic test. Diagnostic tests are used to identify problems or difficulties faced by students, then to plan for following up in the form of solving efforts based on identified problems or difficulties. Diagnostic tests are designed to detect the difficulties in learning outcomes of students so that in preparing diagnostic tests must be designed according to the format and response that the diagnostic test. In addition, diagnostic tests are developed based on an analysis of the sources of errors or difficulties that might be the cause of problems.

The researchers gain information from one of the schools in Yogyakarta, namely "Indonesia" Yogyakarta Vocational High School related to the set of final semester assessment tests at the school that had not been analyzed so that the quality of the questions tested by students was unknown. To find out the characteristics of this test given to students are valid and reliable can be seen from the item analysis. The purpose of this analysis is to know which items are good and which items are not good, then the teacher can look for possible reasons why the items are not good, the good or bad of these items are seen from the difficulty level of the questions, item discrimination, and the effectiveness of answer choice [5]. This is in accordance with Retnawati [6] that in order to obtain a valid and reliable assessment instrument, the teacher must analyze the items which aim to study and examine each item so that the quality of items are obtained before the items are used, item analysis includes qualitative analysis (content validity and construct) and quantitative analysis (level of difficulty, item discrimination, validity, and reliability of item). This is in line with the statement Retnawati [7] that the instrument to be used must be analyzed to show the instrument with good qualifications. This quality includes the validity and reliability as well as the characteristics of the items arranged in the instrument.

After the analysis, it is expected that indicators can be identified that have difficulty levels such as high, medium, and easy levels. If the indicator is categorized as difficult, it can be a reflection of the teacher in order to make learning strategies that help students to understand difficult material. Difficult test items can be identified based on classical test theory. A program that can be used to help to conduct item analysis is the Quest program. This program is able to analyze dichotomy data and polytomous data.

Based on the explanation above, there is a need for research on the characteristics of the set of final semester test to determine the validity, reliability, difficulty level, differential power, and effectiveness of the distractor on the items for final semester test in even semester grade XI academic year 2018/2019 in “Indonesia” Yogyakarta Vocational High School.

2. Method

This research was a descriptive explorative with a quantitative approach. Descriptive exploratory research was a research method by obtaining data in the form of number analysis according to the method used then interpreted [8]. In this study, the object of the research was the final semester test set
and the student answer sheet obtained from the final semester of even semester grade XI mathematics subject academic year 2018/2019 in "Indonesia" Yogyakarta Vocational High School. The data collection technique in this study was a secondary data source through the documentation on the set and the student answer sheet then makes the item indicator. The final semester test set that has been obtained and has been made item indicator was validated by 3 experts judgment. The score from this validation was used to prove the validity of the instrument based on the Aiken index.

Based on data collection, the test set consists of one package there were 30 items of multiple choice. The students who answer the test were 229 students. To obtain an instrument that has high quality, it is necessary to analyze the items empirically. The data obtained is analyzed for its characteristics based on classical test theory. Data analysis techniques used in this study are validity, reliability, level of difficulty, item discrimination, and effectiveness of distractors. Data in the form of a final semester test set and student answer sheets were analyzed with the help of the Quest program.

Validity is the level of accuracy of an instrument to measure something that must be measured, with other tests being said to be valid if the test measures what is to be measured [9]. The content validity is measured based on the Aiken index. This Aiken index is a rater agreement index for the suitability of items with indicators that want to be measured using the item [10]. The validity criteria are if less than 0.4 is stated low, if the value greater than or equal to 0.4 and less than or equal to 0.8 is stated medium, and if greater than 0.8 is stated high category. The reliability of the instrument is stated in the internal consistency of the Quest program result. The difficulty level of the items is categorized by looking at the criteria according to Gooding [11]. The difficulty level criteria are if \( p \) greater than 0.7 is stated easy category, is if \( p \) greater than or equal to 0.3 and less than or equal to 0.7 is stated medium category and if \( p \) less than 0.3 is stated high category. The item discrimination is categorized by looking at the criteria referring to Retnawati, Arlinwibowo, and Sulistyaningsih [10]. Item discrimination index said goof category when the value greater than or equal to 0.3, if the value less than 0.3 then it is not good category. Item distractor effectivity is categorized as an effective distractor if the value greater than or equal to 5%, if the value less than 5% the it is not an effective distractor.

3. Result
The results of the study describe the validity, reliability, level of difficulty, item discrimination, and effectiveness of the distractor on the items in the final semester test for mathematics subject grade XI 2018/2019 in "Indonesia" Yogyakarta Vocational High School. The following is a description of the results of this study.

While the test set and student answer sheets obtained, the researcher arranges a blueprint. This blueprint includes three-dimensional material, vectors, application of exponential equations, social arithmetics, geometrics transformations. The item indicator that compiled is validated by expert judgments. This validation is done to determine the validity of the instrument. Furthermore, Retnawati, Arlinwibowo, and Sulistyaningsih [10] explain that validity will show support for empirical facts and theoretical reasons for the interpretation of test scores or scores for an instrument, and is related to the accuracy of the measurement. In this study, the content validation was determined using expert agreement. The agreement of this expert can be known as the item validity index based on Aiken’s index. Based on the Aiken index and the specified criteria, the content validity of the instrument is obtained as follows.

| Category   | Number of items                          |
|------------|------------------------------------------|
| Low        | -                                        |
| Medium     | 11                                       |
| High       | 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30 |
Table 1 shows that the validity of the final semester test for even semester grade XI mathematics subject academic year 2018/2019 in "Indonesia" Yogyakarta Vocational High School is 29 valid items in the high category and 1 valid item is in the medium category.

In addition, to prove validity, reliability estimates are also carried out. The reliability of an instrument is the consistency of the instrument if it were given to the same subject even though the people are different, the time is different, or the place is different then it will give the same or not significantly different results [9]. The reliability estimation is done with the help of the Quest program with the term internal consistency, the instrument reliability coefficient is 0.77. The reliability coefficient of this instrument is at a high correlation and can be interpreted as having good reliability.

Based on calculations with the Quest program, it can be seen the difficulty level of the item. According to Jamal [12], item difficulty is defined as the percentage or proportion of test-takers (students) who correctly answer the item. The statement can be interpreted that the level of difficulty is the percentage or proportion of test-takers who answer correctly. Jabbarifar [3] states that difficulty items indicate the percentage of students who respond correctly to a test item. Item difficulty is expressed on a scale from 0.00 to 1.00. A value of 0.00 indicates that no students are responded correctly; a value of 1.00 indicates that all students are responded correctly.

To solve the item number 30, the step that can be done is to convert the borrowed money from the converted year in 1.5 years × 12 months = 18 months. After that, determining the interest for 18 months 1.2% × 18 = 21.6%. Interest during the borrowing of money is known to further determine the amount of money that must be paid during the payment of money 18 months \[
\frac{121.6\%}{100} \times 15,000,000 = 18,240,000.
\]

Then, it can be determined that the amount of money for 18 months is 18,240,000 - 15,000,000 = 3,240,000. So, the interest that must be paid after 1.5 years is Rp. 3,240,000.00.

Based on Table 2, it is known that the items with easy difficulty level are 4 items. A total of 21 items were included in the difficulty level in the medium category. Item questions with difficult difficulty levels have 4 items. Furthermore, the results of the analysis of the items are based on item discrimination.

**Table 2. Item of difficulty level category**

| Category | The number of items | Items |
|----------|---------------------|-------|
| Easy     | 4                   | 16, 24, 26, 27 |
| Medium   | 21                  | 1, 2, 3, 4, 6, 7, 9, 10, 11, 12, 13, 14, 17, 18, 19, 20, 22, 23, 25, 28, 29 |
| High     | 4                   | 5, 8, 15, 21 |
|          |                     | 29     |

Aby borrowed IDR 15,000,000 in a bank with a single rate of 1.2% per month. The interest that must be paid after 1.5 years is ...

A. Rp 1,980,000.00
B. Rp 2,000,000.00
C. Rp 2,100,000.00
D. Rp 2,120,000.00
E. Rp 2,160,000.00
answers obtained are not in the answer choices provided. Because the answers are not in the correct answer choices, when analyzing classical test theory data with the help of the Quest program, item number 30 is not included.

Based on the analyzed data, it was seen from the proportions to answer correctly, it could be obtained data items that were difficult for students. In this package of final evaluation questions, out of 30 items, there are 4 items which are in the difficult category, namely numbers 5, 8, 15, and 21. The four difficult items on average proportion answer correctly less than 30%. The most difficult problem is number 15 indicated by as many as 37 students answered correctly from a total of 229 students. The level of difficulty or the proportion that answers correctly is 0.162 or the percentage that answers correctly is 16.2% and the item discrimination is 0.10. Here is the question of number 5.

Given vector \( \vec{a} = (-6, -8) \) and vector \( \vec{b} = (-3, 4) \). If the angle size between the two vectors is 60°, then the result of \( |\vec{a} + \vec{b}| \) is ….

A. \( \sqrt{7} \)  B. \( 2\sqrt{7} \)  C. \( 3\sqrt{7} \)  D. \( 4\sqrt{7} \)  E. \( 5\sqrt{7} \)

To solve the questions in question number 15, students must know the resultant vector, namely the length of the vector from the sum of the vectors, in this problem there are two vectors so that it is formulated as \( |\vec{a} + \vec{b}|^2 = |a|^2 + |b|^2 + 2|a||b|\cos\theta \). After knowing the formula, students can determine the length of the vector \( a \) and vector length \( b \).

\[
|a|^2 = \left(\sqrt{(-6)^2 + (-8)^2}\right)^2 = (\sqrt{36 + 64})^2 = \sqrt{100}^2 = 100
\]

then we obtain \( |a| = 10 \). Next step,

\[
|b|^2 = \left(\sqrt{(-3)^2 + (4)^2}\right)^2 = (\sqrt{9 + 16})^2 = \sqrt{25}^2 = 25
\]

we obtain \( |b| = 5 \). After that, substitute the value of \( |a|^2, |b|^2, |a|, |b|, \) and \( \cos 60^\circ \) on the formula so that it is obtained \( |a + b|^2 = 100 + 25 + 2 \cdot 10 \cdot 5 \cos 60^\circ \). Then doing the ordinary algebraic operation is obtained \( |a + b|^2 = 175 \), then the shape is changed to \( |a + b| = \sqrt{175} \). Obtained the answer is \( |a + b| = 5\sqrt{7} \). Based on interviews with mathematics teacher, students faced the difficult to understand the concept of the resultant material of this vector. Students find it difficult to associate the resultant vector with known vector values, consequently, students cannot find which formula to use. In addition, students also have difficulty understanding the information provided so that there are students who immediately add up vectors \( \vec{a} \) and vectors \( \vec{b} \) regardless of the angle between two vectors. The concept that must be mastered by students in order to solve this problem is the length of the vector, the angle between the two vectors, and the addition of vectors.

Another difficult item is number 21 indicated by as many as 68 students correct from a total of 229 students, the level of difficulty or the proportion that answers correctly is 0.298 or the percentage that answers correctly is 29.8% Discriminant power of 0.36. Here is the question of number 21.

The image of rotation 90° about the origin \( O(0,0) \) a line equation \( 2x - y + 3 = 0 \) is . . .

A. \( 2x - y - 3 = 0 \)  B. \( 2x + y + 3 = 0 \)  C. \( x + 2y + 3 = 0 \)  D. \( x - 2y + 3 = 0 \)  E. \( x - 2y - 3 = 0 \)

To solve the problem, it is necessary to know if the point is (x, y) is rotated 90° about the origin O(0, 0), then the image is (y, -x)

\[
2x - y + 3 = 0
\]

\[
2y' - (-x') + 3 = 0
\]

\[
2y' + x' + 3 = 0
\]

\[
x' + 2y' + 3 = 0
\]

So, the image of the equation is \( x + 2y + 3 = 0 \). In this material, the difficulties for students are in substituting point shadows on line equations. Students can understand the point transformation but have not mastered the line transformation. Students at least master the concept of point rotation, line equations, and determining the equation of line shadows.

Item discrimination refers to how good an item can discriminate or differentiate among test takers who differ on the construct being measured by the test [11].
Table 3. Item discrimination criteria

| Criteria | Category | The number of the item |
|----------|----------|------------------------|
| ≥ 0.3    | Good     | 21                     |
| < 0.3    | Not Good | 8                      |
| Total    |          | 29                     |

Table 3 shows that there are 21 items included in the good category and 8 items included in the not good category. It can be interpreted that as many as 21 items were able to discriminate well students who mastered competencies based on certain criteria.

Next, look at the effectiveness of the distractor. The effectiveness of this distractor aims to see how many students choose each answer option that has been provided. A good distractor, if it has a value of more than 0.05 on each answer option, means that more than 5% of students choose the correct choice of answers and the deception that has been provided.

Table 4. Item distractor effectivity

| Criteria | Category       | The number of the item |
|----------|----------------|------------------------|
| ≥ 5%     | Effective      | 29                     |
| < 5%     | Not effective  | 0                      |
| Total    |                | 29                     |

Based on Table 4, there are 29 items that had effective distractors while there were no items with distractors that were not effective distractors, ie less than 5% of students answered the questions.

4. Discussion

The validity of the final semester test is 29 valid questions in the high category and 1 valid item is in the medium category. The valid instrument indicates that the test set is able to measure the students' ability or measure what should be measured. Its statement in line with Allen and Yen [13] that the validity of a set test can be interpreted as the ability of a test to measure what should be measured.

The reliability estimation is done with the help of the Quest program with the term internal consistency, the instrument reliability coefficient is 0.77, it means a high consistency and can be interpreted as having good reliability. Ebel and Frisbie [14] state that if the test has high consistency, the test is accurate, reproducible, and generalizable to the same opportunity for testing and instruments.

Based on the analysis, there are four items included in the high difficulty level category. Difficult material covering vector material determines the result $| \vec{a} + \vec{b} |$ determines the number of squares of the roots, the third dimension determines the length of the projection, and the transformation determines the shadow of the line. The four topics included in the difficult category, based on the information of the teacher who teaches in class XI of "Indonesia" Yogyakarta Vocational High School, the difficulties faced by students in understanding concepts, connecting concepts, errors in applying procedures, and selecting or seeking information in order to obtain answers right. The cause of the difficulty is in line with previous studies conducted by Ali [15] that the difficulties of students in doing the problem one of them are a weak understanding of the concept. The statement is also supported by Bradley, Notar, Herring, and Eady [16] that students' weaknesses in understanding are closely related to the mathematical concepts used to solve problems. The same thing was expressed by Jazuli, Setyosari, Sulthon, and Kuswandhi [17] that the difficulties of students include a lack of understanding of students in understanding concepts and often using the formula incorrectly in
solving problems. In addition, students' weaknesses in basic mathematical abilities are understanding concepts in line with statements of Rittle-Johnson, Schneider, and Star [18] and Prismana, Kusmayadi, and Pramudya [19].

The existence of the difficulties described above, such as the lack of understanding of students' concepts, needs improvement or improvement in the quality of learning. According to the teacher's statement, in order for students able to understand the concept they must review the material that has been learned then learn the next material and provide a lot of exercises for students to get used to solving problems. This is in line with the research conducted by [20] that learning with models in which there is repetition, namely repetition which means deepening, expansion, stabilization by the way students are trained through assignments or quizzes can improve students' mathematical understanding. The lack of students' ability to understand concepts can also be supported in the learning process that is applying a context that is close to everyday life. This statement is in line with the statement of Ali [15]. This is also supported by Booth, Lange, Koedinger, and Newton [21] that understanding the concept can be deepened by providing training to students through activities related to real objects that are around. Karimah, Kusmayadi, and Pramudya [22] also argues that the relationship between concepts and symbols is needed with things or objects that have been known by students. Furthermore, Tambychik, Meerah, and Aziz [23] state that learning strategies that significantly influence the understanding of concepts are contextual learning strategies.

The ability to understand concepts can affect their procedural abilities. This is in line with the research conducted by Nebesniak [24] that conceptual knowledge competencies and procedural knowledge support each other in increasing the second competency of this knowledge. In this study, in addition to students having difficulty understanding the concept, students also experience the ability to complete the procedure. Difficulties in implementing this procedure are in line with the research of McCowan and McCowan [25] as many as 72.6% of students have difficulty in implementing procedures in solving problems. In relation to the ability to apply the procedure, according to the teacher, students can be trained in procedural fluency with exercises that are interesting, varied, and achieve the goal of mastering certain things. This is in line with the statement Arifin [26] that the exercises provided are brief, interesting, purposeful, and distributed. In practice, teachers can apply appropriate techniques to help facilitate the procedural fluency of students. According to Booth, Lange, Koedinger, and Newton [27] worked examples can be given to students as learning tools and allow teachers to analyze student work procedures.

Choosing or searching for the information needed is also a problem for students, this is related to the ability to understand information and understand images. The statement relates to research that has been done by Karimah, Kusmayadi, and Pramudya [28] that students experience difficulties at the highest level, namely language ability, information, and visual-spatial. This difficulty is also like the results of research [29] which found that the weaknesses of some students were the ability to understand numbers, visual-spatial, and information. The important thing that affect of item characteristics is language related to word problem or mathematical expression. This statement in line with Kan and Bulut [31] that linguistic complexity levels are able affect the mathematics item such as word problem and mathematical languages.

Based on this study, students encounter difficulties in certain material and students got low achievement or achievement is not optimal is caused not only because of the ability of internal students but also influenced by the teacher's performance during teaching. This is in line with previous research conducted by Jailani and Retnowati [31] which states that the teachers 'burden causes less optimal learning. Not much different from the research conducted by Domino [32] that the habits of students in classroom learning have a large influence on students' attitudes to the subject. Thus it is necessary to apply various student-centered learning models so that students in the class play an active role in each learning to improve understanding of students' mathematical concepts. Nebesniak [24] further states that effective learning helps students whether teachers need to plan lessons that focus on three functions, namely how to teach mathematics that can develop concepts, make circular connections, and attract the attention of students to learn mathematics.
There are 21 items included in the good category and 8 items included in the not good category. It can be interpreted that as many as 21 items were able to discriminate well students who mastered competencies based on certain criteria. High performers should be more likely to answer a good item correctly, and low performers more likely to answer incorrectly in other words that it is able to indicate that the item discriminates in favor of the upper group which should answer more items correctly [33].

There are 29 items that had good distractors while there were no items with distractors that were not good, i.e. less than 5% of students answered the questions. Distractors are said to function effectively if the choices were chosen by the test participants from the lower group, whereas if the choices were chosen by the upper group then the distractor is not functioning properly.

5. Conclusion and Suggestion
The test set was valid and reliable. There is one item that has no key so no analysis is carried out. There are 21 items which are medium difficulty items, 4 items with low difficulty levels while 4 items are difficult. Items that are felt difficult by students need to be indicated. For further research, this kind of research is recommended but analyzed based on item response theory. In addition to math subjects, research can also be done on other subjects. The next recommendation is to create items test in proportional difficulty, regard to language or linguistics levels, and examine the teacher's strategy in teaching mathematics so that students do not experience significant difficulties.

References

[1] Biggs J 1999 *Int. J. Phytoremediation* 21 57–75
[2] Hadi and Elvira 2016 *J. Eval. Pendidik.* 4 58–68
[3] Jabbarifar T 2009 *Proc. 2nd Int. Conf. Teach. Learn. (ICTL 2009)* 1–9
[4] Sridadi 2002 *Int. J. Heal. Sci. Res.* 8 26–37
[5] Direktorat Pembinaan SMA 2010 Juknis Analisis Butir Soal di SMA
[6] Retnawati H 2017 *Olympiads in Informatics* 11 123–42
[7] Retnawati H 2016 *Heri Retnawati* 9 786021 547984
[8] Ebel R L and Frisbie D A 1991 *Essentials of Educational Measurement*
[9] Retnawati H, Kartowagiran B, Arlinwibowo J, and Sulistyaningsih E 2017 *Int. J. Instr.* 10 257–76
[10] Retnawati H, Arlinwibowo J, and Sulistyaningsih E 2017 *Int. J. New Trends Educ. Their Implic.* 8 28–38
[11] Gooding S 2009 Children’s Difficulties with Mathematical Word Problems. Sara Gooding 29 31–6
[12] Jamal F 2014 *J. MAJU* (Jurnal Pendidik. Mat. 1 18–36
[13] Ismail S F Z H, Shahrill M and Mundia L 2015 *Procedia - Soc. Behav. Sci.* 186 474–81
[14] Purnomo B 2018 *Soulmath* 6 1
[15] Ali T 2012 *SAGE Open* 2 1–13
[16] Bradley J R, Notar C E, Herring D F and Eady C K 2009 *Asian Soc. Sci.* 4 60–5
[17] Jazuli A, Setyosari P, Sulthon and Kuswandi D 2017 *Glob. J. Eng. Educ.* 19 49–53
[18] Rittle-Johnson B, Schneider M and Star J R 2015 *Educ. Psychol. Rev.* 27 587–97
[19] Prismana R D E, Kusmayadi T A and Pramudya I 2018 *J. Phys. Conf. Ser.* 1008
[20] Rohrer D 2009 *J. Res. Math. Educ.* 40 4–17
[21] Booth J L, Lange K E, Koedinger K R & Newton K J 2013 *Learn. Instr.* 25 24–34
[22] Karimah R K N, Kusmayadi T A and Pramudya I 2018 *J. Phys. Conf. Ser.* 1108
[23] Tambychik T, Meera T S M and Aziz Z 2010 *Procedia - Soc. Behav. Sci.* 7 171–80
[24] Nebesniak A L 2017 *Math. Teach.* 106 354–8
[25] McCowan R J and McCowan S C 1999 *Res. Found. SUNY*
[26] Arifin Z 2017 *Theorems* 2 28–36
[27] Booth, J. L., Lange, K. E., Koedinger, K. R., & Newton, K J 2013 Learning and Instruction 25 24–34.
[28] Karimah, RKN., Kusmayadi, TA., & Pramudya, I 2018 J. Phys Conf. Ser.
[29] Booth, J. L., Lange, K. E., Koedinger, K. R., & Newton, K J 2013 Learning and Instruction 25 24–34.
[30] Tambychik T, Meerah T S M & Aziz Z 2010 Procedia Social and Behavioral Science 7(C) 142-15.
[31] Jailani J & Retnawati H 2016 The Online Journal of Counseling and Education 5 1-13.
[32] Kan A and Bulut O 2017 *J. Educ. Hum. Dev.* 4 133–46
[33] Domino, J 2009 Teachers’ influence on students’ attitudes toward mathematics. *Research and Teaching in Development Education* 26 32-54
[34] McCowan, R.J & McCowan, S.C 1999 Research Foundation of SUNY Center for Development of Human Services 20 1-39