Learning achievement of Elementary School student of mathematics using the Testlet model instrument: A comparison between the 2006 Curriculum and the 2013 Curriculum

S Hamdi and B Kartowagiran
Universitas Negeri Yogyakarta, Yogyakarta, Indonesia

E-mail: syukrulhamdi@uny.ac.id

Abstract. This study aims to describe and compare elementary school students’ achievement in mathematics using the 2006 curriculum and 2013 curriculum with a testlet model. The research used a survey method and a quantitative approach. The participants were 552 from 14 elementary schools consisting of 273 students from elementary schools applying the KTSP curriculum and 279 from elementary schools applying the 2013 curriculum. The instruments used a testlet model test instruments that had met valid and reliable criteria. The data were analyzed using independent sample t-test. The results of the analysis show that there were not significant differences in the two curricula with t value with equal to 1.246 on the significance level of 0.118. The data obtained using a testlet model were more detailed related to students’ difficulties, students' ability to answer questions and the level of guessing from students. The results can be used as a basic in enhancing student mathematics learning.

1. Introduction
Curriculum is an inseparable item of the education process, the intended curriculum that applies in Indonesia is the 2013 curriculum. The implementation of the 2013 curriculum has not been entirely carried out evidenced by the large number of schools in the regions which remain to implement the KTSP curriculum (in Bahasa: Kurikulum Tingkat Satuan Pendidikan or called as the 2006 curriculum). Several interesting distinctions are studied associating with the 2013 curriculum and the KTSP curriculum in particular with the subject of mathematics since they relate to how they are presented at the elementary school level. In KTSP, mathematics stands as one special subject while mathematics in the 2013 curriculum becomes a set of unit with other subjects namely Science, Social Sciences, Indonesian Language, and the related subjects called integrative thematic.

Based on the research finding on the implementation of the 2013 curriculum [1], it concluded that the quality of mathematics learning assessment was in the poor category. One factor of the problem is the absence of a theoretical and empirical problem analysis process. Consequently, mathematics subjects are quite often included in the National Middle School / High School National Exam or National Standard School Examination for Primary Schools. This shows that mathematics is an essential subject to be mastered by students, one of which is to solve various problems in other lessons such as physics. it has been reinforced by the [2] revealing that "students is obligated to and experience that mathematics is a language tool that is able to translate wide range of problems, and
help us overcome problems in various cases". It means that the mathematical abilities possessed by students can be benefited to cope with multifarious problems and difficulties obtained in learning a wide kind of knowledge and insights, particularly those related to science.

De Lange’s [3] divided the objectives of mathematical education into three levels covering the goals of lower level, medium level, and high level. In addition, De ‘Lange’s presents levels of understanding in a form of a pyramid like Figure 1. The picture shows that at the lower level, the aim of mathematics education is that students are able to master mathematical concepts in line with objects of knowledge, definitions, technical skills, and algorithms standards such as sum of integers and/ or fractions.

Figure 1. Three levels of mathematical education objectives

At the second level, which is the moderate level, the purpose of mathematics education is marked by students' understanding of two or more concepts, making relationships, integration, and problem solving which are terms that are often used at this level. At the third level, named the high level, students' understanding is characterized by their ability to work with complex materials such as mathematical thinking and reasoning, communicative, critical, creative, interpretive, reflexive, generalizing, and mathematical concept. Haylock and Thangata [4] suggested "problem solving is when the individual use think mathematical knowledge and reasoning to close the gap between the givens and the goal" that problem solving occurs when a person thinks about mathematical insights and reasons a case to close the gap between reality what happened and what was expected. Hence, solving problems should require creativity to think scientifically and make logical reasoning as well as use realistic problems.

At the ground base, a problem is said to be realistic if the problem can be imagined in students' mind and does frequently a problem exist in real life [5]. It indicates that if the ability of students to imagine a solution taken in accordance with a previously owned-knowledge scheme will significantly help learners solve wide ranges of mathematical problem which they found and become a part of realistic mathematics at the same time. If this happens, it will indirectly affect the students' ability to imagine solutions in their real life with logical and systematic thinking for other problem life.
The recognition of a mathematics understanding derived from relational problems can merely be realized maximally if mathematics learning is undertaken optimally and sustainably. One of the efforts that can be done to achieve it is teachers' innovation in developing the materials and instruments which are so obviously needed by students such as a form of story question. Based on data from the Mathematical PPPPTK Training Need Assessment (TNA) Mathematics [6], it turned out that story question is a problem for teachers in teaching and students in learning. It is the causes why they are encountering trouble getting information regarding learning the development of a story question items and its examples. Using the test instrument of testlet model is expected to become an alternative used to answer students' basic needs for increasing their understanding as well as their mathematical knowledge. Yue & Hong-Yun [7] revealed that the testlet model comprised a set of shared-based-stimulus multiple choice items. Therefore, it is based on realistic problems. It is also strengthened by Paek, et al, [8] who stated that in recent years, the intended model has evolved to accommodate those that have the structure of a testlet model where the grain of items has a shared-stimulus. In addition, Edward [9] revealed that the design of the testlet model is basically used in education and research field to see the test participant's responses to a question model whose the answer comes from one paragraph.

The opinions of the experts indicate that the assessment using the testlet model is able to be used to measure students' learning outcomes. The learning result certainly has a high accuracy in order that the learning outcome of those two different curricula can be described and compared clearly. Based on theoretical and empirical studies conducted by Purwo Susongko [10] turned out that in terms of scoring, the testlet model is more practical than the description because scoring can be done objectively and is polytomous. Therefore, the use of the testlet model is very useful to show the standardized comparison of each response given by the test taker. In other words, the form of the testlet model is quite simple while the information provided is quite good.

2. Method
The research method stands for design, participant, data collection and data analysis.

2.1. Design
The survey research was used on this study with an approach of mix method meaning that quantitative and qualitative. Quantitative approach was used to describe the mathematics learning outcome of fractional materials for elementary schools applying the 2013 curriculum and elementary schools applying the KTSP curriculum. The learning outcome gained was a grounded comparison of results between curricula applied. Qualitative approach was used to know the deep factors which distinguished mathematics learning outcomes for fractional material in two curricula implemented in East Lombok and even throughout Indonesia.

2.2. Participant
The population of this study were all elementary school students in East Lombok regency, West Nusa Tenggara province, Indonesia. The sampling technique used was Classified Random Sampling, which a sample was taken randomly based on the geographical location school of the district, border/city/village/mountain. The number of samples used were 552 taken from 14 elementary schools consisting of 273 students from elementary schools applying the KTSP curriculum (1 district city school, 3 border schools, 3 village / mountain schools) and 279 from elementary schools implementing the 2013 curriculum (2 district city schools, 3 border schools, and 2 rural / mountain schools).

2.3. Data collection & data analysis
Data collection used a mathematical test instrument of testlet model, interview guidelines and assessment sheets for experts. Test instrument was tested and examined by seven experts, the results of expert assessments were analyzed using the CVR and CVI formulas. The instrument trial resulted were analyzed using the Item Response Theory using the Parscale program and the reliability
estimation of the instrument used Cronbach Alpha. Differences in mathematics learning outcomes for the KTSP curriculum and the 2013 curriculum were analyzed using independent sample-t-tests.

3. Results and discussion

The results of analysis and discussion were divided into two parts, namely 1) the results and discussion of the quality of the instrument and 2) the description and comparison of students’ learning outcomes implementing the 2013 curriculum and the 2006 curriculum.

3.1. The results of the instrument quality analysis

The analysis was to know the validity index point using Content Validity Ratio (CVR), while it was also used to notice the validity content point, which counted by CVI (Content Validity Index) statistic used to see the index of test content validity. The estimation showed that the KD accordance CVI value and indicator were 0.94 and 0.88, respectively, meaning that the two-instrument grille developed, was valid. The value of CVI accordance point towards indicator achieved 0.97 and 0.89 respectively which meant that the two instrument packages of mathematic testlet model developed were valid based on the experts. Next, the mathematic test instrument reliability for 2006 curriculum was 0.822 and instrument test of mathematic testlet model test for 2013 curriculum was 0.808. This result showed that the test instrument in mathematic testlet model test for class assessment in elementary schools (SD/MI) implementing 2006 and 2013 curriculum were reliable.

Having analyzed the validity and reliability estimation, it was continued with the analysis of the question quality using item response theory (IRT). Analysis using IRT should meet several prerequisite analyzes or test assumptions including unidimension, local independence and invariance parameters. One way to check the unidimensional was to conduct factor analysis resulting in eigenvalue. According to Hambeleton, Swaminathan, and Rogers [11] who assumed that unidimension was able to be presented in eigenvalue plot showing a dominant component. If the eigenvalue of a factor had more dominant value compared to another eigenvalue factor, it could be said that the unidention criteria was fulfilled. The result of 2006 curriculum (package 1) and 2013 curriculum (package 2) test are presented in the following figure 2.

![Figure 2](https://example.com/figure2.jpg)

**Figure 2.** Graphic of Eigenvalue for questions in package 1 and 2

The two eigenvalue graphics above shows that there are only one dominant component seen from 1 curaman both from respondents in package 1 and 2. According to Hambeleton and Swaminathan [12] the existence of a dominant dimension could be reported that the assumption of unidimention was fulfilled.
The second assumption was local independency used to know to what extent the connection between the participants’ answer and each question presented. Local independency assumption fulfilled when the participants are not influenced by the other participants in this testlet test model. According to De Mars [13], local independency could be detected by proving the unidimentional assumption. The local independency could also be seen based on Chi-Square ($\chi^2$) result, the result of $\chi^2$ achieved $\chi^2$counting was smallest compared to $\chi^2$ Table, meaning that local independency was fulfilled.

The next assumption that should be fulfilled was invariance parameter. Invariance parameter was used to know the characteristic of each question that was not depending on students’ ability parameter and the vice versa. According to Hambleton, Swiminathan and Rogers [11] invariance parameter could be checked by using the difference test level for the test participants group.

Invariance assumption towards the ability in this research was proved by dividing the participants into two; answer towards odd numbers and even numbers. Referring to Hambleton, Swiminathan and Rogers [11] who said that “Invariance of ability parameters over randomly equivalent forms (e.g. ability estimates based on examinee performance on the odd-numbered items and the even-numbered items) indicates the variability due to the sampling of test items”. The estimation result from odd and even numbers would be designed scatter-plot, line $y=x$ which would get the information about invariance parameter ability. The ability test result is presented in the following figure 3.

**Figure 3. Scare-plot assumption of invariance ability of participants from package 1 and 2**

Based on the figure 4 above, it can be concluded that invariance assumption is fulfilled because there is a tendency within the spots which are close to $y=x$. The correlation result between the ability of odd participants and even 2006 curriculum (package 1) is 0.981 (high category). Based on the correlation result above, it can be concluded that invariance assumption is fulfilled, so do with the invariance assumption in 2013 curriculum (package 2) it is achieved that the correlation between the ability of odd and even participants is 0.642 (middle category). Base on this correlation, it is clear that the invariance assumption is fulfilled.

In general, the characteristic of mathematic testlet model test for the 2013 curriculum which has been developed with -0.642 to 2.668; different power from 0.432 to 0.941. The characteristic of mathematic testlet model for 2016 curriculum which has been developed has -0,441 to 3,603 level of difficulties; defirrence power 0,312 to 0,808. The selection of 2PL GRM in the research based on the conformity model test result. In general the ability of Elementary School students goes to middle category which exists between -2 to +2. The scoring guide used is fixed depending on the assessment of experts and book guide of mathematic testlet test model because it completed with some ways and example directly.
3.2. The Description and comparison of students’ learning outcomes between the 2013 curriculum and 2006 (KTSP) curriculum

The student’ learning outcomes applying the 2013 Curriculum and the 2006 Curriculum were not much different, it is illustrated from the average value obtained in Table.1

| No | Description | 2013 Curriculum | 2006 Curriculum |
|----|-------------|-----------------|-----------------|
| 1  | Mean        | 28,857          | 30,582          |
| 2  | Minimum     | 0,00            | 2,22            |
| 3  | Maximum     | 93,33           | 84,44           |

Table 1 shows that the rate score of students who learn mathematics using the 2006 curriculum is higher than those using the 2013 curriculum. However, after being tested using an independent sample t-test, it can be concluded that there is no significant difference with a t-value of 1,246 on the significance level of 0.118. It proves that despite having different averages, there are no statistically significant differences. In other words, the learning outcome of students who implements the 2013 curriculum and the 2006 curriculum do not affect student mathematics learning outcomes. This analysis is obtained based on students’ learning outcomes in the same items. In Table 1, it seems that the minimum and maximum scores have a fairly high range, especially in the 2013 curriculum. It is marked that students who are tested using the test instrument testlet model cannot solely answer by guessing as a reason that if item 3 is true but item 1 and 2 are false in one group of the testlet, then it counts as 0. This is caused the testlet used are tiered in 1 item group, students may not answer correctly number 2 if number 1 is not answered correctly as well as number 3 in 1 testlet group. The frequency of student scores is as follows.

![Figure 4](image-url)  

**Figure 4. The frequency of students’ learning outcome**

Figure 4 illustrates that score gained more than 75% respondents obtains score under 40 of 100 in scale. This result is noticeably concerning and needs to be improved. Several items categorized as difficult for students and the results of confirmation to the teachers related to students' answers will be presented randomly. The testlet point number 7 simplifying the fraction also goes to difficult category since the students has lack of understanding in simplifying the fraction. Based on the interview result with an elementary school in Sakra said that:
“most of students are just able to simplify using a simple method if the counter and odd denominator so the counter and the denominator will be devided into 2 and it is considered as the simplest while it can be simplify more”

The result of analysis using Parscale for the testlet point number 7 simplifying the fraction resulting 75.5% or 206 test participants. The students were tricked with a choice A \( \frac{12}{14} \) because \( \frac{24}{28} \div \frac{2}{2} = \frac{12}{14} \), this result can be simplify to \( \frac{12}{14} \div \frac{2}{2} = \frac{6}{7} \). The result of students for testlet point number 14 (point 40, 41 and 42) is 52.4% with wrong answer, 27.1% with right answer and only one question out of their testlet 12.5% the students are able to answer only two question perfectly out of three testlet, 5.5% gave a right answer for testlet point and 2.6% could not give any answer. This question goes to hard question since the students are not able to understand the question provided and the concept of decimal number combination to the close unit is not fully understood. Based on the interview result with one the teacher in elementary /SDN 5 Suralaga with Z initial about the difficulties in testlet point number 14:

“most of students already understand the concept of how to round off the decimal towards the closest unit if there is one or two decimal after coma, but when it is more than two place after coma, they will have difficulties to determine which number should be rounded off, for example in the question to round off the figure 24,325 many students give answer with 24,32”

The question number 15 could be put as difficult question category since many students are not able to change the mix fraction to simple fraction form. This example could be seen in the changing form question \( 6 \frac{1}{2} \) to the simple form is tikk getting the difficulties so that impacting the students’ answer. Base on the interview with the teacher in SDN 2 Aikmel, SDN 5 Suralaga, and SDN 3 Selong that are implying 2013 curriculum, they stated that “the fourth grade of elementary school are still not be able to do mix fraction question” however, 16.3% students are able to give right answer to the mix fraction question. Based on the confirmation to the participants after doing the task, the students who are able to answer the question about mix fraction claimed that they can answer the question since they get course/private outside the school either from parent or course teacher.

The same case also exist in the instrument test of mathematic testlet model test who are using 2013 curriculum, there are 3 questions has the ideal difficulty level in number 4, 5 and 10. The group testlet poin number 4 the question is hard. This statement is based on the analysis result using parscale which is showing that only 2.2% who got the 3 score and 57.7% gave a wrong answer or got 0 score. The students have difficulties due to the same reason with the question number 15 in the instrument using 2006 curriculum where the students are getting difficulties to change the form mix fraction to the simple fraction.

Furthermore, for the group of testlet point number 5 can be categorized to unfit and not good question since the difficulty level is too hard. Based on the discussion result with expert, it is clear that for the question number 5 where the fraction as the comparation is too high for the fourth grade students of elementary school (SD) even for the fifth grade they might not be able to finish the task. To finish those questions, the students should make a fraction then make the fraction to be decimal, however, the expert does not asking to be omitted so the researcher is still able to use and put the question as part of 15 group of testlet point. After doing and widen the test it goes to unfit category and then the researcher omitted since the question is not able to be used for measuring what should be measured.
The last question can be categorized as poor item for the instrument using 2013 curriculum is group testlet point number 10 (point 28, point 29 and point 30). The analysis result using parscale is 73.5% where all students give wrong answer, 13.3% where the students are able to answer point 28 only, 7.9% are able to answer point 28 and 29 with only 3.6% are able to give right answer or achieve a perfect remark within the group of testlet point number 10. This question is involving in difficult question since the students do not comprehend the characteristic of two-dimensional figure especially in regular hexagone in the question. Based on the observation result and asking question directly to the students in SDN 3 Selong, SDN 3 Pancor and SDN 3 Rempung, it is found that the rotating symmetry has been discussed but it is forgotten since they are not focus on the subject and only thematic lesson and the characteristic of two-dimensional figure. The students statement are also supported by the teachers, the fourth grade teachers in those three school said “teacher does not give the mathematic material deeply since it goes to thematic lesson”. Presentation of the results above is the result of student learning that was tested using a mathematical test instrument model testlet. This shows that the testlet can be a solution in the assessment. These results are reinforced [14] explaining that teacher assessment on testlet model mathematics test instrument for classroom assessment was good which showed that the testlet could be a solution for teachers in assessment.

4. Conclusion

The improvement of the curriculum was consistently being encouraged by the government including the origin of the 2013 curriculum which was expected to enhance the 2006 curriculum. Changes in the curriculum are supposed to be able to advance the students' competence regarding mathematics learning. Found on the results of the study, there were no differences to the students' learning outcomes from schools implemented the 2013 curriculum compared to students coming from schools implemented the 2006 curriculum. Even the result of description showed that the rate of students' learning outcome applying the 2013 curriculum were lower than those implementing the 2006 curriculum, although the statistic showed no significance. This was why learners in the 2013 curriculum did not focus on learning mathematics, yet integrative thematic, on the other hand, the 2006 curriculum focused on learning mathematics. The absence of significant differences should be the government's attention to continue to study and review curriculum policies for the mathematics lesson in order that the change of curriculum are also balanced by raising the quality of students' learning outcomes.

References

[1] Abrory, M. & Kartowagiran, B. 2014. Evaluasi implementasi kurikulum 2013 pada pembelajaran matematika SMP negeri kelas VII di kabupaten sleman. Jurnal Penelitian dan Evaluasi Pendidikan, 2 (1), 50-59.
[2] Humenberger, H.1997. Applicable mathematics education teaching mathematics. Diambil pada tanggal 25 Juli 2014 dari http://www.ex.ax.uk/cimt/ijmtl/hhempres.pdf).
[3] Van den Heuvel, M.-Panhuizen. (1996). Assessment and realistic mathematics education. Utrecht: CD-β Press.
[4] Haylock, D., & Thangata, F.2007. Key concept in teaching primary Mathematics. London: SAGE Publication.
[5] Wijaya, A. 2012. Pendidikan matematika realistik: suatu alternatif pendekatan pembelajaran matematika. Yogyakarta: Graha Ilmu.
[6] Rahardjo, M & Waluyati, A. 2011. Modul matematika SD program bermutu pembelajaran soal cerita operasi hitung campuran di Sekolah Dasar. Yogyakarta: PPPPTK Matematika Kementrian Pendidikan Nasional.
[7] Yue, L. & Hong-Yun, H. 2012. When should we use testlet model? A comparison study of Bayesian testlet random-effects model and standard 2-PL Bayesian model. Journal Acta Psychologica Sinica , 44 (2), 263−275.
[8] Paek, I, et al., 2009. Random parameter structure and the testlet model: extension of the rasch testlet model. *Journal of applied measurement*. 10(4), 394-407.

[9] Edward H. Ip. 2010. Interpretation of the three-parameter testlet response model and information function. *Applied Psychological Measurement* 34(7) 467–482.

[10] Susongko, P. 2010. Perbandingan keefektifan bentuk tes uraian dan testlet dengan penerapan graded response model (GRM). *Jurnal Penelitian dan Evaluasi Pendidikan Tahun 14, Nomor 2, 2010*.

[11] Hambleton, R.K., Swaminathan, H., & Rongers, H.J. 1991. *Fundamental of item response theory*. Newbury Park, CA: Sage Publication Inc.

[12] Hambleton, R.K., dan Swaminathan, H., 1986. *Item response theory: Principles and application*. Boston, MA: Kluwer-Nijhoff.

[13] DeMars, C. 2010. *Item response theory, understanding statistics measurement*. New York. NY: Oxford University Press.

[14] Hamdi, S., Kartowagiran, B., & Haryanto, H. 2018. Developing a testlet model for mathematics at elementary level. *International Journal of Instruction, 11*(3), 375–390. https://doi.org/10.12973/iji.2018.11326a.