Higher Order Thinking Skills Students Through Multi-Representation Test on Newtons Law Study

H Z Puspitaningrum*, Wasis, and T Prastowo
Science Education Study Program, Postgraduate Program, Universitas Negeri Surabaya, Surabaya 60231, Indonesia

*Email: hidayah.18030@mhs.unesa.ac.id

Abstract. This study aims to describe higher-order thinking skills in Newton’s law study based on a multi-representation test. The methods used quantitative descriptive and implemented on one of the class X High School in Surabaya academic year 2019/2020. The questions that tested were verbal, graphic, visual, or mathematically representations consist of the ability to analyse, evaluate, and create adjusted cognitive domain in Bloom’s taxonomy revised. Validity and reliability instrument test in this study with help Winstep software. Based on the result of the analysis, it was found high order thinking skill student in the high category of 42 %, high order thinking skill student in the middle category of 16 %, and high order thinking skill student in the low category of 42 %. Higher-order thinking skills had categorized as enough. Average value person logit equals -0.95 which means that more students answered incorrectly on the questions were tested. Student’s ability in representing physics knowledge shows how the student’s level of understanding.

1. Introduction
Higher-order thinking skills are important skills in the 21st century. Therefore, students are required to have the ability to think critically, solve problems, creativity, communicate, and collaborate. Indonesia in this 21st century is facing many global challenges, students need thinking, verbal communication, cooperation, creativity, research skills, and problem-solving to compete in the future. Thinking can connect from one concept to another with a series of thinking, speaking, reading, writing, seeing, listening, and counting [1]. Educations in Indonesia still needs strengthening in using higher-order thinking skills as an international standard recognized by Indonesia. So that students who have good higher-order thinking skills will be able to apply the knowledge and skills that have been developed and can use them to solve new problems through the concepts that have been taught.

Higher-order thinking skills according to Anderson and Krathwohl in Bloom’s Taxonomy revised are cognitive ways of thinking in higher-order thinking levels consisting of levels of analyse (C4), evaluate (C5), and create (C6) [2]. Testing student solving and problem-solving can be used to view higher-order thinking profiles. Provides the results of activities in obtaining the information needed to make decisions about students. For that, we need to test questions that provide a high level of thinking.

Multi-representation ability can help students in understanding concepts, with understanding these concepts students can improve their thinking skills. Mastery of concepts with the use of various scientific language, such as verbal, visual, symbols, and equations that allows students to learn physics [3]. The concept of representation in the form of the verbal sentence, sign as mathematical symbols, pictures, and graphics which can be used as data collection information [4]. In physics lessons, students are expected to have multi-representation abilities. Multi-representations are important to use...
in physics learning, especially in helping to understand concepts and convey scientific ideas [5]. Students who have the ability to represent will have a good chance of problem-solving effectively [6-8].

Studying physics, not only synonymous with mathematical formulations and symbols, but also physics studies verbal, visual, and graphic explanations. In mastering the field of physics, besides having to understand concepts, students must also have the correct mathematical, logical, and intuitive skills [9, 10]. Studying physics will be easy to understand if it is presented in a different representation format, especially on Newton’s law material. The ability of students to explain physics concepts in different representations will show the quality level of students’ thinking abilities.

Research using multiple representations has been done a lot, but it is still rare to do multiple representations on diagnostic tests, especially on higher-order thinking skills. This research was conducted to provide information and identify higher-order thinking skills in Newton’s Law material on the multi-representation test. This research is expected to contribute to education, the results of the assessment in the form of students’ higher-order thinking abilities.

2. Method

The research method used descriptive quantitative data. The research design model used one-shot case study, design model can see in Figure 1. The sample in this research is 12 students of SMAN 17 Surabaya academic year 2019/2020. The sampling technique is cluster random sampling. The quantitative data collection in this research data obtained through a multi-representation test. Multi-representation tests based on the indicator of conceptual understanding in Bloom’s theory analyse (C4), evaluate (C5), and create (C6) [2]. The quantitative analysis of the test instrument includes measuring the level of question validity, and reliability. Descriptive research aims to get higher-order thinking skill students. Data analysis by Winstep software to determine the general description and answer the problems properly.

![Figure 1. Research Design](image)

Information:
- X : Learning activities by teacher
- O1: Implementation used multi-representation tests

3. Result and Discussion

This study described high-order thinking skills based on the result of a multi-representation test on Newton’s law study using the Rasch Winstep software version 4.5.1. Analysis using the Rasch model was used to analysis the appropriate rating scale as a means of assessment for students in teaching [11-13]. Items were analysis include of item validity, item reliability, item difficulty level, item suitability, and item bias.

3.1. Analysis of Multi-Representation Test to Measure Higher Order Thinking Skills

Multi-representation is a form of a concept that describes the same concept with different formats to describe a situation or a system. The test questions had been testing have multiple representations that vary consisting of verbal, visual, graphic, and mathematical.

3.1.1. Indicator of Test Questions. The multi-representation question tested to measure higher-order thinking skills consists of 3 essay questions with reference to higher-order thinking skills according to the revised Bloom's taxonomy which consists of analyse, evaluate, and create. Indicators of test can be presented in Table 1.
Table 1. HOT Indicator and Multi-Representation Shape

| Indicator HOT | Multi-representation |
|---------------|----------------------|
| 1. Analyse the acceleration base on the force exert accord to the concept of Newton's law that applies in everyday life | Visual, verbal, and mathematical |
| 2. Assess the two conditions present due to the friction force and the resulting pull force | Verbal, graphic |
| 3. Plot the maximum angle on the slope of a plane that has a different statistical coefficient of friction | Visual, mathematical, and verbal |

The questions were tested in the analyse (C4) category, students were required to analysis problems in everyday life related to Newton's law material. Representations on the test of analyse (C4) category were a) visual representations, students were able to draw diagrams correctly, b) verbal representations, students were able to provide input based on available choices, and c) mathematical representations were used as supporting evidence to completed. The questions were tested in the evaluate (C5) category students were asked to give judgment base on the problems presented. Representations on the evaluate (C5) category were a) verbal representation, students give answered with their easy understand own sentences to give conclusions supported by data and give evidence that supports their language skills, b) graphical representations that students were able to complete the graphs, and their components such as titles, axis, and data. Representations on the test of create (C6) category were a) visual representations, described free diagrams according to questions that were related correctly, b) mathematical representations, solving mathematical calculation problems, and c) verbal representations that students must develop their own language sentences for interpreting the results of a mathematical answer that had done.

3.1.2. Reliability and Validity of the Instrument. Analysis of the test instruments used, the overall response pattern, and the interaction between person and items can be seen in statistic summary on Table 2 based on tests using Winstep software.

Table 2. Statistic Summary

| Total Score | Count | Measure | Infit MNSQ | ZSTD | Outfit MNSQ | ZSTD |
|-------------|-------|---------|------------|------|-------------|------|
| Mean        | 8.0   | 3.0     | -0.95      | 0.89 | -0.14       | 0.99 | -0.05 |
| S.D         | 3.0   | 0.0     | 2.21       | 0.71 | 1.00        | 0.81 | 1.04 |
| MAX         | 12.0  | 3.0     | 2.23       | 2.35 | 1.30        | 2.53 | 1.41 |
| MIN         | 4.0   | 3.0     | -4.13      | 0.10 | -1.38       | 0.09 | -1.39 |

Real RMSE 1.08  True SD 1.82  Separation 1.69  Person Reliability 0.74
Model RMSE 0.95  True SD 1.89  Separation 1.99  Person Reliability 0.80
S.E. Of person Mean = 0.64

Person RAW Score-to-measure correlation = 1.00
Cronbach alpha (KR-20) Person Raw Score “Test” Reliability = 0.80  SEM = 1.28

Data analysis using Winstep version 4.5.1 software. The result of the Cronbach alpha (KR-20) showed of 0.80 which means that between students and items show good quality or have a high degree of reliability. In the measure column and the mean row on Table 2, this logit shows the average score students of question items had been giving. The mean value obtained was -0.95, this value was smaller
than the logit value of 0.00 which means that students’ ability is smaller than the difficulty level of the questions.

Reliability person showed the consistency of the person to items. In Table 2, the reliability of the person obtained a value of 0.74 which means that consistency of the person to items is a fair. So, this test can be concluded that the multi-representation test is reliable to use. Because valid calculations are reliable, but reliable calculations are not yet valid [14, 15].

Validity analysis or item suitability analysis level can be used to qualify the following conditions
a) The value Outfit Mean Square (MNSQ) received is $0.5 < \text{MNSQ} < 1.5$,
b) The Outfit Z-Standard (ZSTD) received is $-2 < \text{ZSTD} < +2$,
c) The value Point Measure Correlation (Pt Mean Corr) $0.4 < \text{Pt Measure Corr} < 0.85$.

Analysis of item suitability are shown in Table 3.

| Item | Total Score | Total Count | Measure | Outfit | Ptmeasur | Al | Item |
|------|-------------|-------------|---------|--------|----------|----|------|
|      |             |             | MNSQa   | ZSTDb  | CORRc    | Exp|       |
| 28   | 12          | 0.85        | 1.14    | 0.45   | 0.83     | 0.83| E1    |
| 35   | 12          | -0.63       | 0.97    | 0.07   | 0.85     | 0.83| E2    |
| 33   | 12          | -0.21       | 0.86    | -0.19  | 0.84     | 0.83| E3    |

a MNSQ: Mean Square  
b ZSTD: Z-Standard  
c CORR: point measure correlation

According to Bond and Fox, the value of the outfit means square, outfit z-standard, and point measure correlation are the criteria used to see the level of item fit. If the items on the three criteria are not suitable, it means that the item is not good so that needs to be repaired or replaced [16]. In Table 3, the questions tested suitable the level of suitability of the items or was valid.

3.2. Analysis of Student Abilities

The measured ability was high-order thinking skills based on the results of the multi-representation test. Higher-order thinking skills are the thinking skills of individuals with complex mindsets to describe the material, construct representations, analysis, and make conclusions.

3.2.1. Pearson Measure. Based on the analysis results using Winstep software, a summary of the data is obtained in Table 4. Table 4 gives information about each student's logit. Students with codes L17 and P27 had the highest logit score, which was +2.23, it was indicates that the students had the highest score or answered the questions correctly. Meanwhile, students with code L22 had the lowest logit is -4.13, which means that the student had the lowest score or answered questions with many errors. The student's average score was -0.95, this value was smaller than or less than 0.00, so it can be analysis that the proclivity of students to answer this multi-representation test questions is much answer wrong.

To find map information about the difficulty level of the questions on each item and the students' ability can see in Figure 2. Figure 2 showed distribution of students’ abilities on the left and distribution of the difficulty levels of the items on the right. From figure 2, it can be analysis that on the left map there are three students (P17, P24, and P14) who had a high level of ability from all the difficulty levels of the questions were given. So that the three students get the best score than the others. Two students (L20 and P23) who can did two essay questions well, namely questions E2 and E3, but cannot solve correctly for E1 items that had a higher difficulty level. Then five students (P18, P15, L13, L21, and P19) were unable answered the questions correctly on the questions with the lowest items problem difficulty, namely E2. This means that these students need special attention in solving multi-representation questions to measure higher-order thinking skills.
If we look at the picture above, the distribution of students’ abilities was not in the same place, it showed that the abilities of 12 students were different, and the questions given can provide the information needed to determined students' higher-order thinking skills.

Table 4. Pearson Measure

| Entry number | Total Score | Count | Measure | Infit MNSQ | Infit ZSTD | Outfit MNSQ | Outfit ZSTD | Person |
|--------------|-------------|-------|---------|------------|------------|-------------|------------|--------|
| 5            | 12          | 3     | 2.23    | 2.35       | 1.30       | 2.53        | 1.41       | P17    |
| 12           | 12          | 3     | 2.23    | 0.10       | -1.23      | 0.09        | -1.27      | P24    |
| 2            | 11          | 3     | 1.13    | 1.19       | 0.51       | 1.32        | 0.63       | P14    |
| 8            | 10          | 3     | 0.31    | 0.51       | -0.52      | 0.58        | -0.36      | L20    |
| 11           | 10          | 3     | 0.31    | 0.94       | 0.17       | 0.86        | 0.07       | L23    |
| 4            | 9           | 3     | -0.37   | 0.22       | -1.38      | 0.22        | -1.39      | L16    |
| 6            | 7           | 3     | -1.61   | 1.81       | 1.12       | 1.80        | 1.10       | P18    |
| 3            | 6           | 3     | -2.28   | 0.95       | 0.20       | 0.98        | 0.23       | P15    |
| 1            | 5           | 3     | -3.07   | 0.19       | -1.34      | 0.23        | -1.23      | L13    |
| 7            | 5           | 3     | -3.07   | 0.19       | -1.34      | 0.23        | -1.23      | P19    |
| 9            | 5           | 3     | -3.07   | 0.83       | 0.05       | 0.92        | 0.16       | L21    |
| 10           | 4           | 3     | -4.13   | 1.47       | 0.76       | 2.13        | 1.21       | L22    |
| Mean         | 8           | 3     | -0.95   | 0.89       | -0.1       | 0.99        | -0.1       |        |
| P.SD         | 2.9         | 0     | 2.12    | 0.68       | 1.0        | 0.78        | 1.0        |        |

Figure 2. Map Information About the Difficulty Level of the Questions and Students’ Ability
3.3. Accuracy of High Order Thinking Skills on Multi-Representation Newtons Law Test

Higher-order thinking skills are categorized based on the revised Bloom's Taxonomy of analyse (C4), evaluate (C5), and create (C6). The data obtained by test results and analysis using Winstep software, the classification of a person with higher-order thinking skills can be determined using the following Equation:

\[ H = \frac{\text{(4x separation) + 1}}{3} \]  

(1)

With a Pearson separation value of 2.19, so value of \( H = [(4 \times 1.69) + 1] \); \( 3 = 2.58 \). The number 2.58 is rounded to 3, so it can be concluded that students can be divided into three large groups, they are high order thinking skill in the high-level group, the middle-level group, and the low-level group. The percentage of student groups can see in Table 4.

| Category  | Interval Logit                  | Percentage |
|-----------|--------------------------------|------------|
| High      | \( 0.11 \leq P \leq 2.23 \)    | 42 %       |
| Middle    | \( -2.01 \leq P < 0.11 \)      | 16 %       |
| Low       | \( -4.13 \leq P < -2.01 \)     | 42 %       |

Table 4. Percentage of HOTS students

Students' success in solving physics problems needs to be accompanied by understanding using multi-representations. Multi-representation is a very important ability as a basic skill in learning physics. This is in accordance with the opinion that multi-representation serves as a complement to cognitive processes, helps limit errors in other interpretations, and deepens conceptual understanding [17].

Based on Table 4, the percentage of students who had high-order thinking skills in the high-level were 42 %, this is categorized as enough. Based on the results of the multi-representation test to determine students' high thinking skills, students who were high-order thinking skills in the high-level were able to solve problems well. Students can solve problems with more than one representation. This is suitable for the results of research by Kohl which showed that the use of several appropriate representations would help students in physics problems [18]. Students who high-level category are able to solve mathematical, verbal, visual, and graphic problems. They were able to describe the language well so that they were able to provide evidence and supporting evidence well. Students who are categorized as high-level are able to solve problems well, understand the meaning of words, and relate them to physics concepts [19]. Students are able to describe mathematical equations and with visual abilities, these students were able to describe the components of force being tested on Newton's law material.

Based on Table 4, the percentage of students who had high-order thinking skills students' in the middle category was 16 %. Students who categorized as middle-level were almost correct in working on representation problems, but they were not quite right to answer the test. Students who as middle category there were some students who had difficulty in deciphering language according to the correct laws of physics in making conclusions based on the problems presented, students still have difficulty solving problems using representation solving, they think that solving physics problems identical to solving mathematical problems.

Based on Table 4, the percentage of students who had high-order thinking skills students' in the low category was 42 %. Students in this category still need a high accuracy to solve questions. As in drawing graphs, they still had difficulty draw functions on the diagrams presented. Students who are in the category of less skilled in solving problems, students have the habit of using mathematical solutions in solving physics problems, students are trapped in mathematics without understanding the concept of physics [19, 20]. The difficulty that students are they are unable to apply mathematical knowledge in the context of physics [21]. Students with low categorized had not been able to solve the questions that require them to think highly. So that these students still cannot be categorized as
students with high-order thinking skills. Therefore, with this multi representation test, we can group students based on their thinking abilities. A multi-representation instrument can be used to evaluate the same student at different times [22]. Using multiple representations is believed to be able to support students' understanding of depicting one representation to another from concrete to abstract and from simple to complex [23].

Students who not have high-order thinking skills in the high-level category must give special attention so these students are able to take part in learning with even more difficult quality questions. Because at this age, students should have the ability to think abstractly. This is in accordance with the psychological theory presented by Piaget that in children aged 11-15 years and up, children are able to think the concept of reasoning, they can apply logic to hypotheses well in real conditions and involve the use of systematic reasoning. So really needs to be specially trained to students' abilities are equal to students who already have high order thinking skills.

4. Conclusion
The questions that tested were valid and reliable based on the analysis using Winstep software. Based on the results and discussion above, it can be concluded that students' high-level thinking skills in physics are classified as enough. It was evidenced by the results of the acquisition of the percentage of high-order thinking skills students' in the high category of 42 %, high-order thinking skills students' in the middle category of 16 %, and high-order thinking skills students' in the low category of 42 %. Skills are needed in working on high-level questions, especially questions in the form of multiple representations. In this trial, there were still many students who answered incorrectly the questions tested, this was found in the Rasch analysis using Winstep software, it was shown that the student's logit average score was -0.95.

References
[1] Arwood E L 2011 Language Function an Introduction to Pragmatic Assessment and Intervention for Higher Order Thinking and Better Literacy (Philadelphia: Jessica Kingsley Publisher).
[2] Anderson L W, Krathwohl D R, and Airasian P 2001 A Taxonomy for Learning, Teaching, and Assessing (New York: Longman).
[3] Waldrip B 2008 Proc. 2nd Int. Semin. Sci. Educ. (Bandung: Universitas Pendidikan Indonesia)
[4] Hutajulu M, Minarti E D, and Senjayawati E 2019 J. Phys. Conf. Ser. 1315 012023.
[5] Siswanto J, Susantini E, and Jatmiko B 2018 J. Phys. Conf. Ser. 983 012034.
[6] Panaoura A, Michael-Chrissanthou P, Gagatsis A, Elia I, and Philippou A 2017 Int. J. Sci. Math. Educ. 15 723.
[7] Henke A and Hötteecke D 2015 J. Sci. Educ. 24 349.
[8] Sutopo and Waldrip B 2014 Int. J. Sci. Math. Educ. 12 741.
[9] Handika J, Cari, Suparmi, and Sunarno W 2015 Int. Conf. Math. Sci. Educ. 34.
[10] Setyani N D, Cari, Suparmi, and Handika J 2017 Int. J. Sci. Appl. Sci. Conf. Ser. 1 162.
[11] Boone W J, Abell S K, Volkmann M J, Arbaugh F, and Lannin, J K 2011 Int. J. Sci. Math Educ. 9 551.
[12] Lamb R L, Annetta L, Meldrum J, and Vallet D 2012 Inter. J. Sci. Math. Educ. 10 643.
[13] Van Zile-Tamsen C 2017 Res. High. Educ. 58 922.
[14] Ludwig T, Priemer B, and Lewter D 2019 Res. Sci. Educ. 2019 1.
[15] Ursachi G, Horodnic I A, and Zait A 2015 Procedia Econ. Financ. 20 679.
[16] Bond T G and Fox C M 2015 Applying the Rasch Model, Fundamental Measurement in the Human Sciences 3rd Edition (New York: Routledge).
[17] Ainsworth S 2008 Vis.: Theory Pract. Sci. Educ. 3 191.
[18] Kohl P B, Rosengrant D, and Finkelstein N D 2007 Phys. Rev. Spec. Top. – Phys. Educ. Res. 3 010108.
[19] Theasy Y, Wiyanto, and Sujawarata 2018 J. Phys: Conf. Ser. 983 012005.
[20] Kusumawati I, Marwoto P, and Linuwhi S 2015 J. Mater. Sci. Eng. 5 60.
[21] Jannah L N, Sarwanto, and Sukamin 2019 J. Phys: Conf. Ser. 1153 012120.
[22] Samsudin A, Ramalis T, Fratiwi N J, and Aminudin A H 2020 *Int. J. Psychosocial Rehabilitation* 24.

[23] Hartini J, Liliasari S, Setiawan A, and Ramalis T R 2019 *J. Phys.: Conf. Ser.* 1491 012019.