JURNAL PEMBELAJARAN FISIKA

http://jurnal.fkip.unila.ac.id/index.php/JPF

Vol 11 (2), 2023, 81-96 ISSN: 2302-0105 (p); 2684-9828 (e)

Development of High-Order Thinking Skills Assessment Instrument for High School Students on Rotational Dynamics

Siti Musfiroh*, Undang Rosidin, Viyanti

Physics Education Departement, Lampung University, Indonesia

* *e-mail: sitimusfiroh21@gmail.com*

Received: May 26, 2023 *Accepted: March 15, 2023* *Published: March 16, 2023*

Abstract: This study aimed to develop a test instrument for high school physical reasoning abilities related to rotational dynamics. This study aimed to determine: Quality of the Instrument Quality of the Higher Order Thinking Ability Instrument in senior high school Physics, including content validity of the expert judgment, empirical validity based on item suitability using the Rasch and PCM models, reliability, and item difficulty index; and profile of high school physics advanced reasoning test scores. This research is research and development. The development model used is based on the Borg & Gall development model. The research was conducted at MAN 1 Bandar Lampung with 42 students in class XI. Using the Quest program, reliability was analyzed using the Quest program, which was represented by internal consistency values and item difficulty indexes analyzed using the Quest program. The results of the ability measurement profile are grouped into three categories presented in the analysis of the Quest program assessment. Research and development resulted in a high-quality thinking ability test instrument for high school physics. The results are the quality of the tool's content based on the actual assessment in the valid category. Empirical validity is based on the suitability of the item with the Rasch model IRT approach with highly reliable criteria. Task difficulty varies from -2.17 to 1.51. Generally, the results of measuring high-level thinking skills show that students have medium-level skills.

Keywords: Higher Order Thinking Skill, Rotational Dynamics, Rasch Model, Partial Credit Model

DOI: http://dx.doi.org/10.23960/jpf.v11.n2.202302

INTRODUCTION

The 21st century demands changes to the learning process in schools by the demands of qualified abilities so that students can compete globally. Learning is no longer carried out only by the teacher who is in full control, but also by student participation (Tindowen et al., 2017). Students must be involved and active in learning so that they can build their knowledge based on their abilities, while the teacher is only a facilitator (Kim et al., 2015; Eom et al., 2016). In the information era, learning can be interpreted as a learning process carried out by teachers to develop creative students, which can improve students' abilities to build knowledge and control subjects better (Abidin, 2014).

The era of globalization is characterized by advanced humans who are more capable, able to think critically, and able to meet service, quality, and product standards (Bashan & Kordova, 2021). The demands of the globalization era can be anticipated by thinking at a higher level (Ali & Mishra, 2014). Education emphasizes improving performance and quality through learning supported by systems, materials, and resources so that it can produce the best workforce (Primary & Istiyono, 2015). One of the requirements to achieve this is that students must have the ability to think at a high level or HOTS Skills (higher-order thinking Skills). Students must be able to analyze, evaluate, adapt, and be able to apply a constructivist perspective or conceptual process to science (Widhy 2013). Knowledge competency requires students to be curious so that students can understand, apply, and analyze factual, conceptual, procedural, and metacognitive knowledge (Malik et al., 2018).

High-level reasoning skills have become the focus of physics learning (Bao & Koenig, 2019). Students must be trained to work to solve a problem, discover things themselves, and try hard to implement their ideas so that students can understand and apply what they already know (Widana, et al. 2019). The topic of mechanics is one of the physics materials where misunderstandings often occur (Liu & Fang, 2016). One of the mechanics topics that must be mastered in lesson XI in high school is rotational dynamics. It is very important to understand this material because it relates to everyday life (Adawiyah, et al., 2019). However, in reality, many students find it difficult to understand and apply the concept of rotational dynamics (Rahmawati, et al., 2019). Students are unable to analyze and describe free force diagrams and the causes of rotational motion, so they do not understand the problem given (Suparno, 2005). This is related to the method of students' view of problems and requires students to think higher when facing problems in rotational dynamics material (Widana, 2017).

Students must be accustomed to facing questions that not only train their memory but also train high-level reasoning skills. For this reason, the instruments used must be able to measure high-level thinking abilities, where students do not just memorize, repeat, or refer to an answer without processing (Kemendikbud, 2017). In evaluating students' high-level thinking abilities, there are five steps used for measurement, namely transferring one concept to another, processing and applying information, finding relationships in various types of information, using information to solve problems, and reflecting on ideas and information critically and holistically.

It is believed that high-level evaluation skills must be carried out using written test instruments that have been tested and proven, where tests include true/false tests, multiple-choice tests, and multiple-choice tests. (Suryani et al., 2015). Several previous studies have developed PISA model test instruments in the form of multiple-choice (Wibowo & Cholifah, 2018) and a combination of multiple-choice and description (Sinaga, 2015). The multiple-choice question type is more often used in measurements related to objectivity and is easier to apply and process data. However, this type of question cannot reveal students' ability to reason and give reasons (Putri, Istiyono, & Nurcahyanto, 2016). A combination of constructed response tests and multiple-choice tests namely reasoned multiple-choice, is the best alternative (Ku, 2009). However, there has not been much development of test instruments in the form of reasoned multiple-choice questions that reveal HOTS in the PISA model. For this reason, it is necessary to have a high-level thinking ability test instrument that is under the 2013 Curriculum learning to determine students' high-level thinking abilities in Rotational Dynamics material in high school physics.

METHOD

This research is a type of development research. Development was carried out by referring to the Borg & Gall development stage. However, only 7 steps are used in its implementation, including research and data collection, design, product development, preliminary field trials, verification of test results, field trials, and final product improvement. The hypothetical design is shown in Figure 1.

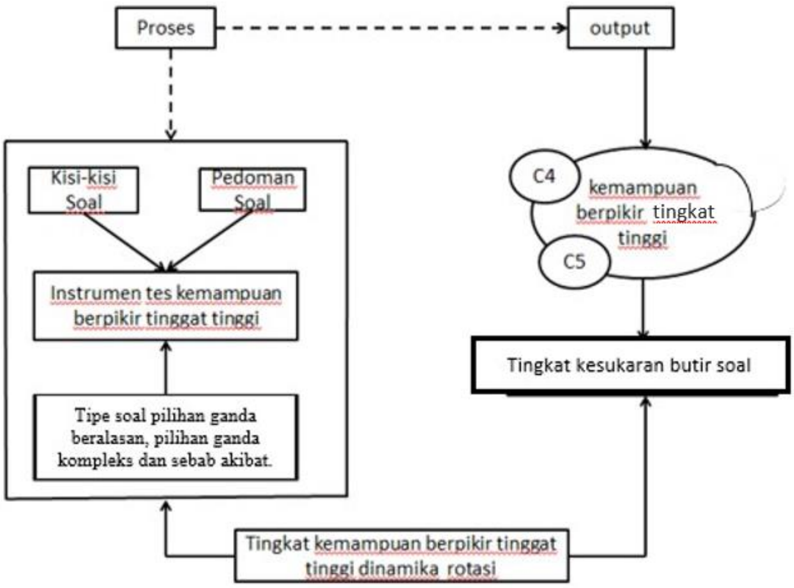


Figure 1. Hypothetical Design

The products developed in this research consist of a reasoned multiple-choice test, a complex multiple-choice test, and a causality test in high school physics subjects using rotational dynamics material to measure students' high-level reasoning abilities. The results of the expert assessment were developed to design educational assessments using HOTS questions according to the modified Bloom's Taxonomy, to measure analytical skills (C4) and (C5). This research aims to develop a high-quality thinking ability test instrument at MAN 1 Bandar Lampung.

The sample used in this research was all students of Class XI MIA MAN 1 Bandar Lampung for the 2022/2023 academic year. The questions used consisted of 15 reasoned multiple-choice questions, 15 complex multiple-choice questions, and 15 causal questions. Questions are scored by giving a weight of 1 to 4 in the form of reasoned multiple-choice questions, while causal and complex multiple-choice questions are given a weight of 1 if the explanation is correct, and 0 if the explanation is correct.

The assessment model used is a polytomous model, where the score given is 0 for a wrong answer, 1 for a correct choice or correct reason, and 2 for a correct choice and reason. Incomplete items were considered missing and were not included in the analysis. The reliability of case estimates or internal consistency states that Cronbach's Alpha value is 0.65 and is included in the sufficient category because it is more than 0.6 (Subali & Suyata, 2011).

An expert review was carried out to obtain the extent of the content validity of the higher-order thinking ability test instrument that had been created. Expert review was carried out by two lecturers at the University of Lampung in the physics education department and one practising physics teacher at the school. Based on an expert assessment of the high-level thinking ability test instrument, an average score of 0.72 was obtained, which is included in the medium category and meets the valid criteria. Analysis of expert assessments based on V Aiken's calculations shows the results in Appendix 3. Thus, the high-level thinking ability test instrument has met the valid criteria and can be tested.

Suggestions for improvements from validators become a reference for researchers to improve the product being developed. Based on expert advice, researchers develop products according to the input for each aspect listed in Table 1.

The assessment scores in the assessment guide have been corrected according to the type of questions shown in Figure 2.

A. Tipe Soal Pilihan Jamak Beralasan

Nomor soal	Skor	Bobot soal
1-15	2	Jika jawaban dan alasan benar
	1	Jika jawaban atau alasan benar
	0	Jika jawaban salah atau soal tidak dijawab

Skor Maksimal Tipe Soal Pilihan Jamak Beralasan: 30

B. Tipe Soal Sebab Akibat

Nomor soal	Skor	Bobot soal
1-15	2	Jika jawaban dan alasan benar
	1	Jika jawaban atau alasan benar
	0	Jika jawaban salah atau soal tidak dijawab

Skor Maksimal Tipe Soal Benar Salah: 30

C. Tipe Soal Memilih Banyak Respon

Nomor soal	Skor	Bobot soal
1-15	2	Jika jawaban dan alasan benar
	1	Jika jawaban atau alasan benar
	0	Jika jawaban salah atau soal tidak dijawab

Skor Maksimal Tipe Soal Memilih Banyak Respon: 30

Untuk menentukan nilai akhir maka:

Jumlah Skor Maksimal : 60

$$\text{Nilai akhir} = \frac{\text{skor yang diperoleh}}{90} x100$$

A. Tipe Soal Pilihan Jamak Beralasan

Nomor Soal	Skor	Bobot soal
1-15	4	Jika jawaban siswa benar, alasan benar dan penyelesaian soal berada pada level 2
	3	Jika jawaban siswa benar, alasan benar dan penyelesaian soal berada pada level 1
	2	Jika jawaban benar, alasan salah
	1	Jika jawaban salah atau alasan benar
	0	Jika jawaban siswa salah dan alasan salah atau soal tidak dijawab

Skor Maksimal Tipe Soal Pilihan Jamak Beralasan: 60

B. Tipe Soal Sebab Akibat

Nomor Soal	Skor	Bobot soal
1-15	1	Jika jawaban benar
	0	Jika jawaban salah atau soal tidak dijawab

Skor Maksimal Tipe Soal Benar Salah: 15

C. Tipe Soal Pilihan Ganda Komplek

Nomor soal	Skor	Bobot soal
1-15	1	Jika jawaban benar
	0	Jika jawaban salah atau soal tidak dijawab

Skor Maksimal Tipe Soal Pilihan Ganda Komplek: 21

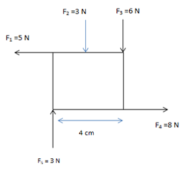
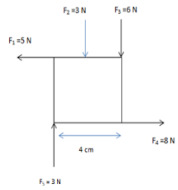
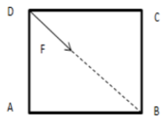
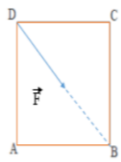
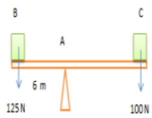
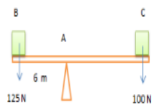
Untuk menentukan nilai akhir maka:

Jumlah Skor Maksimal : 96

$$Nilai\ akhir = \frac{Skor\ yang\ diperoleh}{96}\ x\ 100$$

Figure 2. Fixed Assessment score

Table 1. Suggestions for Improvement by Validators

Validation	Suggestion	Repair Results
Language	Replacing the word square shape with the word square field 2. Terdapat lima buah gaya bekerja pada sebuah persegi seperti gambar dibawah ini. Bangun persegi ABCD panjang setiap sisi 4 cm dengan gaya yang bekerja F_1, F_2, F_3, F_4, F_5. Momen gaya yang terjadi dipusat persegi sebesar… A. 0 Nm B. 3 Nm C. 5 Nm D. 8 Nm E. 10 Nm  **Gambar 2**. Bangun Persegi ABCD	The word square has been replaced with the word square changed the sentence so that it becomes the communicative language 2. Lima buah gaya F1, F2, F3, F4, F5 yang bekerja pada bidang berbentuk persegi seperti gambar disamping. Jika panjang setiap sisi 4 cm momen gaya yang terjadi dipusat persegi sebesar… N.m A. 0 B. 3 C. 5 D. 8 E. 10 
	Improvement on the word build into a field, change the image to a square instead of a rectangle and write a vector sign 4. Terdapat gaya F seperti pada gambar yang berkerja sebesar 10 N pada bangun ABCD dengan sisi sisi 20 √2 cm. Besarnya torsi F dengan poros di titik A adalah … A. 2 cmN B. 2 √2 cmN C. 4 cmN D. 20 √2 cm N E. 200 cmN  **Gambar 4**. Bangun Persegi ABCD	Substituting the word shape into a field, drawing into a square field and writing the correct vector sign 4. Gaya F sebesar 10 N bekerja pada bidang ABCD berbentuk persegi seperti ditunjukkan pada gambar 4, jika panjang sisi bidang tersebut 20√2 cm, maka besarnya torsi F terhadap poros di A adalah … cmN A. 2 B. 2 √2 C. 4 D. 20 √2 E. 200  **Gambar 4**. Bangun Persegi ABCD
Material	Change the weight information to mass and include information on the amount of gravity, just replace the words at the ends 5. Dua balok diletakkan diujung-ujung jungkat jungkit dengan massa yang berbeda balok dititik B = 125 N dan balok dititik C = 100 N. Jika jarak antara titik A dan titik B 6 m, maka besar jarak antara titik A ke titik C.. A. 6 m B. 6,5 m C. 7,5 m D. 8 m E. 9,5 m  **Gambar 5**. Jungkat-Jungkit dengan Beban	Change some of the information in the sentence to make it clearer 5. Dua balok yang massanya berbeda diletakkan diujung jungkat-jungkit seperti pada gambar 5. Balok yang bermassa B = 125 kg diletakkan di titik B sedangkan balok yang bermassa C = 100 kg diletakkan di titik C . Jika jarak antara titik A dan titik B 6 m, maka jarak antara titik A ke titik C adalah… m (g = 10 m/s) A. 6 B. 6,5 C. 7,5 D. 8 E. 9,5  **Gambar 5**. Jungkat-Jungkit dengan Beban
Construct	Improvements and suggestions on causal questions do not use reasons for answers 1. Sebuah silinder pejal mempunyai momen inersia lebih kecil daripada silinder berlubang yang terbuat dari bahan yang sama serta memiliki massa yang sama. **SEBAB** Untuk memberikan percepatan sudut pada sebuah benda berlubang diperlukan lebih banyak tenaga putaran Alasan :	The reason answer is replaced by only filling in the choices in the consequence question 1. Sebuah silinder pejal mempunyai momen inersia lebih kecil daripada silinder berlubang yang terbuat dari bahan yang sama serta memiliki massa yang sama. **SEBAB** Untuk memberikan percepatan sudut pada sebuah benda berlubang diperlukan lebih banyak tenaga putaran Jawaban:

The results of the review using the expert validation test produce items that can be used as valid and reliable questions to measure students' high-level reasoning abilities. Based on an expert assessment of the high-level thinking ability test instrument, an average score of 0.72 was obtained, which is in the medium category and meets the valid criteria. The expert assessment analysis is based on V Aiken`s calculations. Thus, the high-level thinking ability test instrument has met the valid criteria and can be used.

The results of product trials were carried out to obtain the quality of the questions on all items based on the Rasch model. The suitability of the items using the IRT (Item Response Theory) approach and the reliability and difficulty level of the items were examined. The test results were analyzed using the Quest program. The Quest program can analyze dichotomous, polytomous, and combined polytomous data. The results are listed in Table 2.

Table 2.Results of Conformity of Items with the Model

Parameter	Item Estimation	Testee Estimation
MNSQ infit mean and standard deviation	1.00 / 0.24	0.98 / 0.33
The mean and standard deviation infit t	0.1/0.8	-0.03/0.98
Infit MNSQ for each item	0.64 – 1.29	
Infit for each item	-1.5 - 2.0	

Reliability analysis in this research was carried out using the Quests program. The reliability test results presented with internal value consistency have a reliability value of 0.89, which means that the reliability criteria are very reliable according to Guilford (1956). Thus, the advanced thinking skills instrument in high school physics subjects is a reliable measuring tool.

Table 3. Item Difficulty Level Results

No	The Easiest Question Items		The most difficult item	
	Item to	B	Item to	Item to
1	item_32	-2,17	item_35	1,51
2	item_40	-1,76	item_3	1,36
3	item_4	-1,75	item_33	1,00
4	item_20	-1,63	item_29	0,84
5	item_8	-1,48	item_17	0,83
6	item_34	-1,43	item_13	0,81
7	item_22	-1,39	item_9	0,80
8	item_36	-1,15	item_43	0,80
9	item_44	-1,03	item_23	0,77
10	item_24	-1,02	item_25	0,59

Based on Table 3, the difficulty level question which most easily obtained from question 32 with a mark index hardship -2.17. The level hardship question which most difficult is question 35 with a mark index difficulty level of 1.51. The difficulty of question item (b) is based on the greater the bi parameter value, the greater the test taker's ability required to answer the question item correctly. The item is said to be good if the difficulty index is more than -2.0 or less than +2.0 ($-2 < b < 2$) (Hambleton et al, 1991 b). If it is close to -2 then the item is said to be easy, while if it is close to +2 then the item is said to be difficult (Istiyono, 2020).

Level performance students analyzed using program Quests. Evaluate results search analysis gives indication ability participant test. Evaluation skills shared become three-level skills. Results level proficiency student which uses instrument test proficiency thinks tall shown in Table 4.

Table 4. Ability Level of Learners

Category	**Amount**
High	1
Currently	41
Low	0

Based on Table 4, it is known that from a total of 42 students, data was obtained that 1 student had the high ability, 41 students had medium ability and 0 students had poor ability. So based on these data, it can be concluded that more students have average abilities.

RESULT AND DISCUSSION

Results

The quality of the development of the Physics Higher Order Thinking Ability Test Instrument is assessed based on Quest analysis. Empirical validity was obtained which was based on the suitability of the items with the Rasch and PCM models, namely that the average MNSQ infit for item estimation was 1.00 and the standard deviation was 0.24, while the average MNSQ infit was 1.00 and a standard deviation of 0.24. the testee estimate is 0.98 with a standard deviation of 0.33. The average infit t parameter in the item estimate was obtained as 0.1 and a standard deviation of 0.8, while the average infit t in the testee estimate was -0.03 with a standard deviation of 0.98. The INFIT MNSQ value for each item is in the range of 0.64 to 1.29, while the INFIT t value for each item is in the range of -1.5 to 2.0. According to Adams and Kho (1996: 30), an item is said to be fit if the average INFIT MNSQ and INFIT t values in the item estimates and t estimates range between 1 and the standard deviation ranges from 0.0. An item is said to be fit if the INFIT MNSQ value is between 0.77 and 1.33. The goodness of fit of each item can be seen in the INFIT t value, which meets the criteria if the INFIT t value is between -2 to 2. Thus, the value obtained from INFIT MNSQ and INFIT t states that the item is fit to the model. The question items can be said to be valid for carrying out measurements. The results of this research are also by research conducted by Bashoir and Supahar (2018), that the analysis carried out using the Quest program showed that 61 questions had MNSQ infit values in the range of 0.77 to 1.30, so all the items were declared valid.

A suitable measuring instrument for making measurements is a valid and reliable instrument. The reliability of a test instrument shows how far the measurement results are from the accuracy, consistency or stability of a measurement made (Rosidin, 2017: 193). Reliability analysis using the Quest program which is presented on the internal consistency value shows a reliability value of 0.89, which means that the criteria are very reliable according to the reliability criteria according to Guilford

(1956). Thus, the instrument for high-level thinking abilities in high school physics is reliable and can be used to make measurements. Reliability based on internal consistency according to Kuhamaedi (2012) is a method that uses one instrument, so the test is only carried out once. This is because the test can be carried out without the need for repetition so that problems that arise during repetition can be avoided. Calculations using the internal consistency method can also be done using the Alpha Cronbach, Sperm Brown, Kuder, Richadson-20, and Kuder Richadson-21 formulas.

The test results show that the test instrument is valid and reliable. Next, the characteristics of the test instrument regarding the difficulty of the questions were analyzed using the Quest program. Data analysis for dichotomous questions in the form of cause and effect uses the Rasch model and analysis of polytomous data in the form of reasoned multiple-choice questions, and complex multiple-choices using the partial credit model (PCM). This item analysis was carried out because the items had three forms of questions with different scoring. The partial credit model analysis is the result of the development of the Rasch model which obtains analysis results with one parameter, namely the item difficulty index. The analysis results obtained from the items that met the fit criteria with the model had an item difficulty level that ranged from -2.17 to 1.51. An item difficulty index close to -2 is an easy item and an item difficulty index close to 2 is a difficult item. The results of this research are also to the results of research conducted by Abidin et al (2019) that the results of the item difficulty index which had been developed from 136 items were in the range of -0.76 to 0.83 so it was concluded that the item difficulty index for all items is in a good category.

Through the high-level reasoning ability instrument test, the high-level reasoning ability of 42 students can be measured. The results show that there are 1 student with high skills, 41 students with moderate skills, and 0 students with low skills. Based on this data, it is known that the average ability of students is in the middle class with moderate abilities. This shows that the instrument used can measure high-level reasoning abilities.

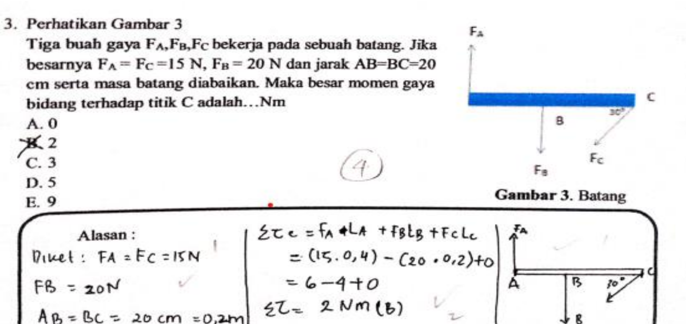


Figure 3. The student's answers to the multiple-choice questions were justified

Figure 3 shows the results of students with correct answers to question number 3. Analysis of student responses shows that they have a good understanding of discourse questions and can process information critically to be able to apply theoretical and mathematical-physical designs to solve the problems given. The three forces acting on the rod determine the magnitude of the moment of the force acting on the flat plane that students are asked about. For a student to get a score of four, the student's reasoning must provide a reasonable response and explanation. This is supported by research by Lewy (2009) which states that students are classified as very good at a high level of reasoning if they can solve questions C4 and C5. For each question answered correctly, students receive four points in this subject. This is in line with research by Fanni et al. (2021) and Intan et al. (2020), where students are classified as competent if they complete (assess) questions at level C5 with a good score and are placed in the quite good category. Based on inquiries "submitted by students", in general students' cognitive abilities can be measured. Samaduri's (2022) research findings show that students' understanding can be determined by the explanations and responses they choose.

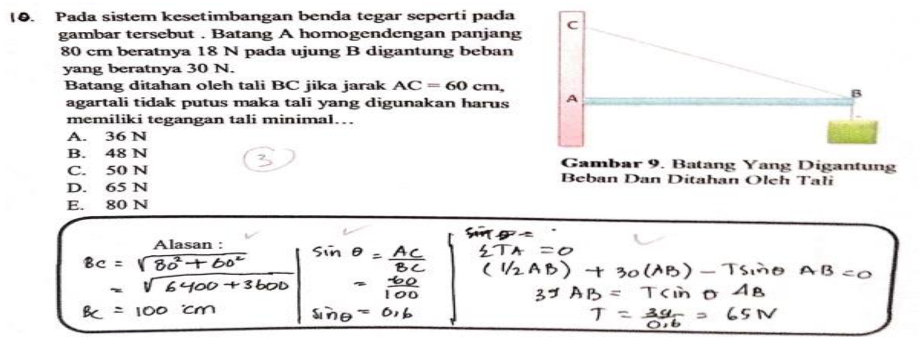


Figure 4. The student's answers to the multiple-choice questions were justified

Figure 4 shows students' correct answers to question number 10 in the form of multiple-choice questions. Students are given information about the assignment. The question is a C6-level question. Based on the student's answer, the student gave the correct answer accompanied by an appropriate explanation, but the score he received was 3. The reason for the student's answer was that the student wrote down the determining angle and determining tension in the rope, but the student did not write specifically beforehand knowing what was known from the question. Lewy's (2019) research findings show that students who excel do so by developing a variety of new problem-solving strategies and creative abilities. Research findings Comparable to the research findings of Wahyuddi et al. (2021), several test takers were able to solve questions at level C6, showing their ability to understand and provide accurate responses.

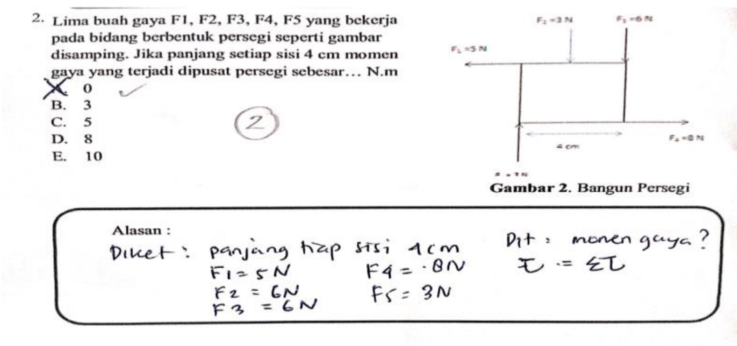


Figure 5. The student's answers to the multiple-choice questions were justified

Figure 5 shows students' correct answers to multiple-choice questions in question number 2. Inquiry This is a C4 cognitive question. Based on the student's answer, option A is correct, but option A does not provide a specific justification for completing the question. therefore the student gets a score of two. Even though they can analyze the question, they are not sure how to write the solution. This is by research by Lewy (2009) which states that "Students' thinking abilities are quite high if students can only solve question C4 and get 2 points for each question answered correctly". This is in line with research by Indra et al. (2019), which revealed that this happened to the majority of students.

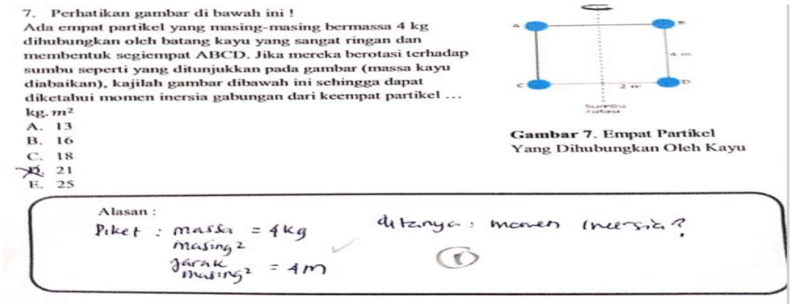
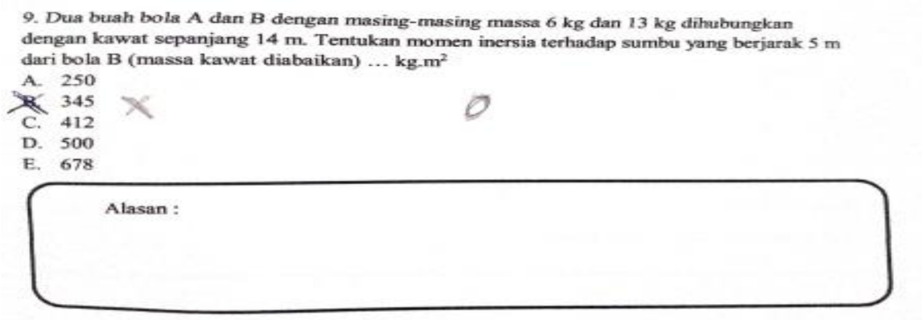


Figure 6. The student's answers to the multiple-choice questions were justified

Based on Figure 6, the student's answer to question number 7 received a score of 1 because he answered incorrectly but gave the correct reason. The task is to provide information about four particles tied into a box with wooden sticks. Students are asked to observe the image to determine the combined moment of inertia of the four particles. This question is a C5-level cognitive question. According to the answer, the student does not have level C5, namely the ability to analyze. Students cannot observe the image of four rectangular particles A B C D, nor can they carry out mathematical calculations to determine the combined moment of inertia. Students' answers show that instead of examining the picture, students write down known information in the question and answer for making wrong decisions. This is by research by Widiyawati et al. (2019), according to him, many students misinterpret images. Students do not master the concept of measurement and usually memorize the formulas given without knowing their meaning. It is difficult for students to use formulas in different contexts. Lewy's research results (2009) show that students can be classified as very good thinkers if they can solve questions C4, C5 and C6, getting 4 points for each question answered correctly and students can solve questions C4 and C5 getting 3 points for each about right. However, the results of students' answers showed that they only got 1 point, indicating that students still had high and low reasoning abilities.

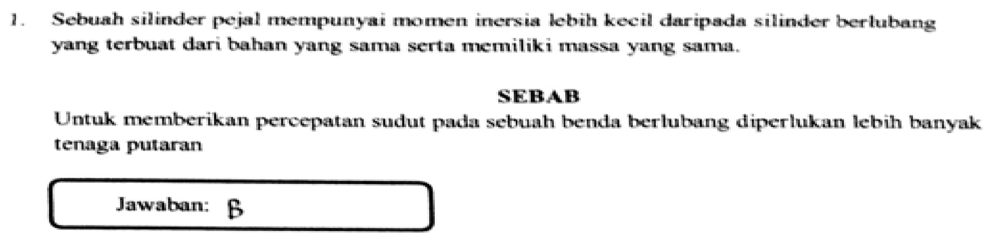

9. Dua buah bola A dan B dengan masing-masing massa 6 kg dan 13 kg dihubungkan dengan kawat sepanjang 14 m. Tentukan momen inersia terhadap sumbu yang berjarak 5 m dari bola B (massa kawat diabaikan) ... kg.m²

A. 250
B. 345
C. 412
D. 500
E. 678

Alasan :

Figure 7. The student's answers to the multiple-choice questions were justified

Figure 7 shows the student's wrong answer to question number 9. In the question, information is given about two balls connected by a string and students are asked to determine the moment of inertia. This question is a C4 cognitive-level question for analysis. Based on the student's answer, it is known that the student's answer is wrong and the student does not give any reason so the score is 0. It is known that the student does not know the method of solving questions at level C4 (analyzing). This is also to the research results of African et al. (2019) that when evaluating the performance level of C4 competency, students are classified as very poor. Due to misinterpretation of questions, conceptual errors and problem-solving strategies, students do not know the method of answering questions well at the analytical level (C4). According to the research results of Saddia et al. (2021) found that the proportion of questions at level C4 was 31.18%, which was included in the low class, indicating that the ability of class XI Science High School students who experienced problems in solving questions at level C4 was still relatively low. Even though the expectations are the same as Wahyuddin's learning results (2021), students who pass the C4 level exam will be able to analyze problems by researching, sorting, identifying and looking for pattern relationship information.


1. Sebuah silinder pejal mempunyai momen inersia lebih kecil daripada silinder berlubang yang terbuat dari bahan yang sama serta memiliki massa yang sama.

SEBAB

Untuk memberikan percepatan sudut pada sebuah benda berlubang diperlukan lebih banyak tenaga putaran

Jawaban: B

Figure 8. Student answers to questions of cause and effect

Figure 8 shows students' answers to questions by consequences. This question is related to level C4. in the question, information is given regarding a problem related to the comparison of the moment of inertia of a hollow cylinder and a solid cylinder made of the same material. Students are asked to differentiate between comparisons. Analysis of students' answers shows that they understand discussion questions well and can process question information critically, enabling students to match true or false statements with a question. Students get a score of 1 because the answer is correct. This is in line with research by Wahyuddin (2021), who believes that students who can complete the C4 level exam can analyze problems by researching, sorting, identifying and finding patterns of information relationships Sari et al. (2019) confirmed the research results. With a total of 50 questions in three variants, consisting of multiple-choice questions, consequence questions, and association questions, the cognitive level is 78% at level C4, 14% at level C5, and 12% at level C6. The average result of class students is 48.17 according to the low criterion. This shows that students still have difficulty reasoning on high-level questions, including questions with formats because of the consequences (RR Sari et al., 2019). Based on the presentation of learning results, it is known that causal questions can be used to measure the reasoning abilities of advanced students at cognitive levels C4 and C5.

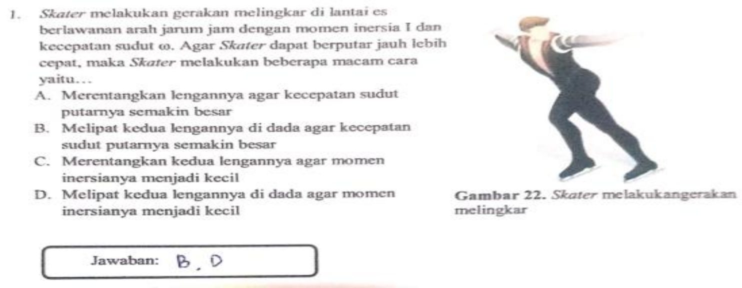


Figure 9. Student answers to complex multiple-choice questions

Figure 9 shows the students' correct answers to complex multiple-choice questions. This question is a question at level C4 (analyzing). Based on the analysis of the answers, students can process information well and critically, so that students can answer correctly in options B and D. Option B, is correct, because the angular speed of rotation is greater, and the step that can be taken is what the athlete can do is fold both arms crosswise. in the chest so that the moment of inertia decreases to a small magnitude. Choice D is correct because when the dancer's arms are extended, the dancer's rotation speed is smaller than when the dancer crosses both arms across his chest. The rotation speed is inversely proportional to the moment of inertia so that the dancer's moment of inertia, when the arms are outstretched, is greater than the moment of inertia when the arms are crossed on the chest. Thus, students get a score of 2 because there are 2 correct answers. Scoring in the form of complex multiple-choice questions means giving 1 score to one correct answer statement. The student's answer is in line with the results of research conducted by Wahyuddin (2021) that students who have been able to take tests at level C4 can analyze a problem by observing, sorting, identifying and finding information relationship patterns.

This is also in line with the results of research conducted by Suseno (2017) which revealed that association multiple-choice, which is one type of complex multiple-choice, influences test takers in working on the questions, which is then supported by an assessment technique that gives weight to each question according to the level. problem difficulty. This influences participants' carefulness in answering questions. The association multiple-choice test is a type of higher-order thinking ability test that trains children to have better thinking abilities than ordinary multiple-choice tests. This is because with the association multiple-choice test form, students will find several choices that need to be analyzed further to answer the question correctly, whereas, with the regular multiple-choice test form, students will have the opportunity to guess the answer. After all, only one answer is correct (Widoyoko, 2019). Based on the presentation of the research results above, shows that complex multiple-choice questions can measure students' high-level thinking abilities.

Discussion

Based on the analysis of students' answers to various types of questions, it is known that students can process the information in the questions critically so that the instruments used can train students' high-level thinking skills. Estimates of students' high-level thinking abilities from the results of trials of high-level thinking ability test instruments show that students have moderate abilities. The results of this research are also similar to those carried out by Istiyono (2017), namely the results of measuring the high-level thinking ability instrument in physics in class .45% medium category, 19.50% high category, and 2.00% very high category. Research conducted by Widiyaningsih (2022) also said the same thing, namely the results of the development of physics questions for high-level thinking abilities based on modern test theory which were designed and presented with Moodle LMS on e-learning which can be accessed online showed that the estimation results show that students' high-level thinking skills are in the medium category.

Students' high-level thinking abilities in the low category are influenced by several factors, one of which is that students are not used to working on high-level thinking ability questions (Tanujaya et al, 2017). They need to be trained to develop high-level thinking skills by being given learning resources based on high-level thinking skills. To realize high-level thinking abilities, students must be required to be more active in learning (Winarti et al, 2015). Teachers are also expected to play the role of facilitators who provide various learning resources and provide feedback on learning (Masruroh & Prasetyo, 2018). Therefore, it is very important to apply high-level thinking skills in learning, so that it can improve and train students' high-level thinking skills.

CONCLUSION

The resulting test instrument is a high-level thinking ability test instrument in high school physics subjects. The physics material tested is odd semester class XI physics material, namely rotational dynamics. The question form consists of reasoned multiple-choice, complex multiple-choice, and cause and effect with a total of 45 questions. The quality of the high school physics high-level thinking ability test instrument has met the test suitability criteria for use in measurement. The content validation of the expert assessment was found to be in the high category and met the valid criteria based on Aiken's V value. Empirical validity based on item fit in the Rasch model for dichotomous data and the PCM model for polytomous data obtained INFIT MNSQ and INFIT t values that fit the model. The reliability estimates presented in the internal consistency values are in the very reliable category with a question difficulty level ranging from -2.17 to 1.51.

Based on the results of measuring high school physics test instruments with cognitive levels C4 and C5, data was obtained that the average level of student's ability to think at a high level was in the medium ability level category.

REFERENCES

Abidin. (2014). Desain Sistem Pembelajaran dalam Konteks Kurikulum 2013. Yogyakarta: PT. Refika Aditama.

Abidin, A. Z., Istiyono, E., Fadilah, N., & Dwandaru, W. S. B. (2019). A computerized adaptive test for measuring the physics critical thinking skills. International Journal of Evaluation and Research in Education, 8(3), 376–383.

Adawiyah, R., Harjono, A., Gunawan, G., & Hermansyah, H. (2019). Interactive e-book of physics to increase students' creative thinking skills on rotational dynamics concepts. In Journal of Physics: Conference Series (Vol. 1153, No. 1, p. 012117). IOP Publishing.

Ali, S., & Mishra, M. (2014). Teaching Strategies for Fostering Higher Order Cognitive Skills: The Need of The Globalised Era. Scholarly Research Journal For Humanity Science and English Language, 1(3), 345-354.

Bao, L., & Koenig, K. (2019). Physics education research for 21st-century learning. Disciplinary and Interdisciplinary Science Education Research, 1(1), 1-12.

Bashan, A., & Kordova, S. (2021). Globalization, quality and systems thinking: Integrating global quality Management and a systems view. Heliyon, 7(2), 1-16.

Deda, N, Y., Ratu, A, H., Amsikan, S., Mamoh, O. (2020). Analisis Kemmapuan Siswa dalam Menyelesaikan Soal Ujian Nasional Matematika SMP/MTs Berdasarkan Perspektif Higher Order Thinking Skills (HOTS). Jurnal Pendidikan Matematika. 3(1).

Eom, S.-J., Youn, J.-J., & Kim, H.-J. 2016. A Study on Image Selection for the Development of Educational Contents Enhancing Undergraduates' Creativity and Personality. Indian Journal of Science and Technology. 9 (26), 1-6.

Fani, K., Fauziana., Rahmiaty. (2021). Analisis Kemampuan Siswa dalam Menyelesaikan Soal HOTS pada Pelajaran IPA Kelas V MIN 25 Aceh Utara. Journal of Primary Education. 2(2), 66-75.

Guilford, J. P. (1956). Fundamental Statistics in Psychology and Education. New York: McGraw-Hill.

Ku, K. Y. (2009). Assessing students' critical thinking performance: Urging for measurements using multi-response format. Thinking skills and creativity, 4(1), 70-76.

Intan, F. M., Kuntarto, E., & Alirmansyah, A. (2020). Kemampuan Siswa dalam Mengerjakan Soal HOTS (Higher Order Thinking Skills) pada Pembelajaran Matematika di Kelas V Sekolah Dasar. JPDI (Jurnal Pendidikan Dasar Indonesia), 5(1), 6-10.

Kim H, Lee H, Youn J, Eom S, & Lee J. 2015. A Study on College Students' Demands for Creativity and Personality Education as Part of the General Education Curriculum. Indian Journal of Science and Technology. 8 (1), 29-36.

Lewy. 2009. Pengembangan Soal Untuk Mengukur Kemampuan Berpikir Tingkat Tinggi Pokok Bahasan Barisan Dan Deret Bilangan Di Kelas IX Akselerasi Smp Xaverius Maria Palembang. Jurnal Pendidikan Matematika. 3(2),14-28.

Liu, G., & Fang, N. (2016). Student misconceptions about force and acceleration in physics and engineering mechanics education. International Journal of Engineering Education, 32(1), 19-29.

Malik, A., Rosidin, U., & Ertikanto, C. (2018). Pengembangan Instrumen Asesmen HOTS Fisika SMA Menggunakan Model Inkuiri Terbimbing. Jurnal Lentera Pendidikan Pusat Penelitian LPPM UM Metro, 3(1), 11-25.

Putri, F. S., Istiyono, E., & Nurcahyanto, E. (2016). Pengembangan instrumen tes keterampilan berfikit kritis dalam bentuk pilihan ganda beralasan (politomus) di DIY. Unnes Physics Education Journal, 5 (2), 76–84

Pratama, S. N., & Istiyono, E. (2015). Studi Pelaksanaan Pembelajaran Fisika Berbasis Higher Order Thinking (HOTS) pada Kelas X di SMA Negeri Kota Yogyakarta. Prosiding Seminar Nasional Fisika Dan Pendidikan Fisika (SNFPF), 104–112.

Rahmawati, I., Sutopo, S., & Zulaikah, S. (2017). Analysis of Students' Difficulties with Rotational Dynamic Topic Based on Resource Theory. Jurnal Pendidikan IPA Indonesia, 6(1). 95-102.

Saddia, A., Sutrisno, S., Saldi, M., & Agriawan, M. N. (2021). Analisis Kemampuan Menyelesaikan Soal Hots Fisika Siswa Sma Di Kota Majene. Phydagogic: Jurnal Fisika Dan Pembelajarannya, 4(1), 1-5.

Samaduri, A. (2022). Analisis Pemahaman Konsep Siswa Yang Diukur Menggunakan Tes Pilihan Ganda Beralasan Pada Mata Pelajaran Biologi. Jurnal Pendidikan Glasser, 6(1), 109-120.

Sari, N., Sunarno, W., & Sarwanto, S. (2018). Analisis Motivasi Belajar Siswa Dalam Pembelajaran Fisika Sekolah Menengah Atas. Jurnal Pendidikan dan Kebudayaan, 3(1), 17-32.

Sari, R. R., Lufri, L., Selaras, G. H., & Darussyamsu, R. (2019). Analisis Kemampuan Berpikir Tingkat Tinggi Peserta Didik Kelas XI SMA Pada Materi Sistem Ekskresi. Bioilmi: Jurnal Pendidikan, 5(2), 91-101.

Sinaga, T. N. (2015). Pengembangan soal model PISA mata pelajaran ilmu pengetahuan alam terpadu konten fisika untuk mengetahui penalaran siswa kelas IX. Jurnal Inovasi Dan Pembelajaran Fisika, 2(2), 194–197.

Suseno, I. (2017). Komparasi karakteristik butir tes pilihan ganda ditinjau dari teori tes klasik. Faktor: Jurnal Ilmiah Kependidikan, 4(1), 1-8.

Setyawarno, D. (2017). Upaya Peningkatan Kualitas Butir Soal dengan Analisis Aplikasi Quest. In Fakultas Matematika dan Ilmu Pengeatahuan Alam, Yogyakarta: Universitas Negeri Yogyakarta.

Subali, B., & Suyata, P. (2011). Panduan analisis data pengukuran pendidikan untuk memperoleh bukti empirik kesahihan menggunakan program Quest. Universitas Negeri Yogyakarta: Lembaga Penelitian dan Pengabdian pada Masayarakat.

Suparno, P. (2013). Miskonsepsi & perubahan konsep dalam pendidikan fisika. Jakarta: Gramedia Widiasarana.

Suryani, A., Siahaan, P., & Achmad, S. (2015). Pengembangan Instrumen Tes untuk Mengukur Keterampilan Proses Sains Siswa SMP pada Materi Gerak. Prosiding Simposium Nasional Inovasi Dan Pembelajaran Sains.

Tindowen, D. J. C., Bassig, J. M., & Cagurangan, J. A. 2017. Twenty-First Century Skills of Alternative Learning System Learners. SAGE Open. 7 (3), 1-8.

Wahyuddin, W., Satriani, S., & Asfar, F. (2021). Analisis Kemampuan Menyelesaikan Soal High Order Thinking Skills Ditinjau Dari Kemampuan Berpikir Logis. AKSIOMA: Jurnal Program Studi Pendidikan Matematika, 10(2), 521-535.

Wibowo, A., & Cholifah, T. N. (2018). Pengembangan instrumen tes tematik terpadu kurikulum 2013 berbasis PISA'S literacy di sekolah dasar. JIPVA (Jurnal Pendidikan IPA Veteran), 2(2), 209–221.

Widana, I. W. (2017). Higher order thinking skills assessment (HOTS). JISAE: Journal of Indonesian Student Assessment and Evaluation, 3(1), 32-44.

Widana, I. W., Parwata, I., & Sukendra, I. K. (2018). Higher order thinking skills assessment towards critical thinking on mathematics lesson. International journal of social sciences and humanities, 2(1), 24-32.

Widhy, P. (2013). Integrative science untuk mewujudkan 21st century skill dalam pembelajaran IPA SMP. Seminar Nasional MIPA, (pp. 1-13).

Widiyawati, Y., Nurwahidah, I., & Sari, D. S. (2019). Pengembangan instrumen integrated science test tipe pilihan ganda beralasan untuk mengukur HOTS peserta didik. Saintifika, 21(2), 1-14.

Winarti, C., Sunarno, W., & Istiyono, E. (2015). Analysis of higher order thinking skills content of physics examinations in madrasah aliyah. In International Conference on Mathematics, Science, and Education 2015 (ICMSE 2015) (Vol. 2015, pp. 32-38).