PJERE, December 2019, Vol. 7, No. 2, pp. 1-10

Profile Higher Order Thinking Skills in Thermodynamics of Senior High School Students

Seftyan Agustihana [1]
Yogyakarta State University, Yogyakarta, Indonesia

Suparno [2]
Yogyakarta State University, Yogyakarta, Indonesia

KEY WORDS

Assessment, Higher order thinking skills, Thermodynamics

ABSTRACT

This study aims to develop a high-level thinking ability instrument of physics subject of senior high school students and measure higher order thinking skills (HOTS) senior high school students of thermodynamic materials in Nganjuk, Indonesia. The study was conducted in three stages, including initial development, trials, and measurements. The initial development stage includes compilation, review, and validation. Validation performed includes content validation by experts. Problems developed in two packages of questions A and B, in which each package consisted of 20 questions, with anchor items as much as 5 items. The measurement results showed that thermodynamic HOTS test kits have obtained empirical evidence fit with partial credit model. The difficulty level of thermodynamic HOTS is in the range of -2.0 to +2.0. The difficulty levels of the most difficult tests in the sequence are aspects of analyzing, evaluating, and creating. The reliability of HOTS has met the high categorical requirements, and based on the total information function of HOTS thermodynamics is appropriately used to measure the HOTS of learners with capability -2.3 to +1.6. The result of the students' response to the ability to analyze is highest, then followed by the ability to evaluate, and create respectively.

Seftyan Agustihana is a Postgraduate Student of Physics Education at Yogyakarta State University, Indonesia.

Introduction

Assessment of education is the process of collecting and processing various information to measure the achievement of learning outcomes of learners, it is stated in the Minister of Education and Culture RI No. 23 of 2016. Collecting and processing of information about the achievement of learning outcomes of learners taken through measurement. Assessment results will be accurate if done using the appropriate instrument. Thus, the assessment of the results of physics learning requires a method of research in accordance with its characteristics; therefore various assessment methods need to be developed.

Educational assessment features are divided into two, including traditional assessment and performance appraisal (Glencoe Science, n.d.). Traditional assessment features include basic knowledge, knowledge processes, content knowledge, and problem-solving. Performance appraisal features include basic knowledge, group learning, self-assessment, application of skills, creative design, authentic applications, creative products, and application of all the skills of learners.

Given the importance of physics in the development of science and technology, requires learners to have Higher Order Thinking Skills (HOTS). So, HOTS assessment is therefore required. HOTS of learners are poorly trained during learning, affecting all aspects of knowledge (Saido, Siraj, Nordin, Bin, & Amedy, 2015). Therefore, the implementation of physics learning should be in accordance with the learning of physics. It is useful to give learners the opportunity to be actively involved in learning science process skills and HOTS.

Components that support HOTS did not develop automatically, but there should be a planned undertaking. Therefore, to improve HOTS, curriculum, teaching materials, learning media and evaluation instruments should be developed with HOTS orientation. Even teachers and students should think with HOTS orientation. The meaning of HOTS automatically involves the ability to remembering, understanding, and implementing (LOTS) (Brookhart & Nitko, 2011). When HOTS develop then LOTS automatically also develop (Richland & Simms, 2015). Not vice versa, HOTS does not automatically develop with LOTS development. For that, teachers and learners need to get used to thinking high level.

Thinking into a skill that can be familiarized. Higher order thinking skills are a conscious access that becomes a logical precondition for controlling a thing that happens (Metcalfe & Dunlosky, 2009). It is not possible to control several processes unless one can consciously think about the activities in this process. Researchers have shown that when individuals are in the process of mastering cognitive activity such as language, memory attention, there is a point where they can perform activities well but cannot consciously contemplate what they are doing (Piaget, 1976). After a successful period of activity, individuals eventually become able to reflect on what they are doing. For Vygotsky, this activity is only higher order thinking when it becomes conscious. Vygotsky explains knowledge in terms of concepts and functions, the acquisition of knowledge is described as a process of internalizing the words and actions of teachers, parents and more on the ability of learners (Kozulin, Gindis, Ageyev, & Miller; 2003). Cognitive processes reflect high-level thinking when they are independent, have social origins, and are accessible to their self-awareness.

Higher order thinking skills are related to the processes it contains. There are several aspects related to higher order thinking skills, namely conceptual understanding, systematic thinking, problem solving, and critical thinking (Brookhart & Nitko, 2011). These aspects can be familiarized and trained through formal education. Education has become one of the means to practice HOTS. Therefore, the process of learning is held with the support of materials, systems, and the best human resources. The development of learning and learning outcomes of learners can be measured using the assessment format.

Assessment techniques of HOTS testing of thermodynamic materials include measurement activities. This is because it can generate numerical data as an attempt to describe the characteristics of learners. The HOTS test of thermodynamic material consists of a number of questions that have a choice of answers and reasons. The test results indicate information that shows the characteristics of learners in the form of HOTS thermodynamic material.

Based on the description above there are several issues that can be raised in connection with HOTS in the learners themselves in physics learning, which can be identified. The learning habit in physics lessons has not fully solved the problem of HOTS, the trained problem of LOTS, so the less developed HOTS.

This study focused on HOTS measurement of thermodynamic subject matter, including the preparation of assessment tools and scaling. Availability of HOTS assessment tools of thermodynamic subject matter can help teachers make HOTS measurements of learners. HOTS measurements were conducted on students of class XI-IPA. Empirical test and measurement is only doing in senior state high schools with assumption that respondent variation is homogeneous. The objectives of this study include, (1) develop eligible devices as HOTS gauges of thermodynamic material, and (2) obtain portraits of HOTS thermodynamic subject matter senior high school students in Nganjuk.

Methodology

Research on the development of HOTS assessment instrument of physics thermodynamic material subjects conducted in the even semester of 2017/2018. The location of research conducted in senior state high schools chosen by purposive sampling

Instrument used in collecting data was written test. The test kit consisted of two package questions: Package A and package B, each package consisting of 20 items of questionable choice. The test kit was prepared by considering the HOTS indicators to be measured.

HOTS test device development model of senior high school thermodynamics subject matter using model modification according to Mardapi (2012). Steps of test development include (1) planning stage, (2) test phase, and (3) measurement phase. The criteria for the preparation of test kits are limited to knowledge with HOTS indicators including the ability to analyzing, evaluating, and creating.

The planning stage begins with goal setting tests. Then compose the test items, and rubrics. The design of test result of development result is then analyzed by expert to get proof of quantity validity of content determined with Content Validity Ratio (CVR) then mean from CVR determine the amount of Content Validity Index (CVI). The elements that experts examine include material, construction, and language aspects.

The test device test phase was conducted to obtain the quantitative test item parameters. The relevant reliability parameters included model matching with PCM, Standard Error of Measurement (SEM) curves, and reliability using Cronbach alpha coefficients. The eligible test items were then assembled into test for measurement.

Measurement in research aimed to estimate the students' ability parameters. Ability estimation results formed the basis for interpretation of the learning outcomes of students as research objects. HOTS measurement results were described quantitatively.

Student response data were analyzed using Partial Credit Model (PCM) model using Quest and PARSCALE program. The Quest program was used to test the suitability of an item with a PCM model. The item is fit if the value of INFIT MNSQ is in the range of 0.7 to 1.30 (Adam & Khoo, 1977). Reliability of the test, the estimate of the difficulty level (the item is declared difficult if difficulty >+2, and very easy if difficulty <-2 (Hambleton & Swaminathan, 1985), and the estimated value of the learner's abilities in logit scale was also measured. The PARSCALE program is used to describe Standard Error Measurement (SEM) parameters.

Findings and Discussions

The HOTS-developed questions with three indicators which include the ability to analyze, evaluate, and create. Matter of thermodynamics material, divided into several submissions taught in high school, was the content covering the Third Law of Thermodynamics, First Law of Thermodynamics, Second Law of Thermodynamics, and the application of the Law of Thermodynamics. The matrix of relationships between HOTS materials and indicators developed, presented in Table 1. In the ability to analyze, divided into two sub-indicators that distinguish and give special characteristics. On evaluating include sub-indicators check and rate. With the ability to create, developed problems with the sub-indicator plan.

Table1

Development of HOTS Problems of Thermodynamic Materials

Sub-material	**HOTS Indicator**		
	Analyze	**Evaluate**	**Create**
Third Law of Thermodynamics	Differentiate the initial temperature and the final temperature in the isobaric process	Selecting factors related to thermodynamic processes	Plan the volume required to produce a desired business
	Determining business by gas due to pressure changes	Assess the great effort made on systems with two different processes	
1st Law of Thermodynamics	Distinguish effort by gas if temperature is different	Choosing the greatest effort on the adiabatic process	Planned negative work on the system
	Determine energy changes in	Assess the volume and final pressure of the gas to isothermal system	
2nd Law of Thermodynamics	Distinguishes the largest and smallest attempts by the system on a heat engine	Check the time it takes for freezing	Design a cooling machine with certain efficiency
	Determine the performance coefficients of the engine efficiency of the heat	Assess entropy changes in melted ice	
Application of the Law of Thermodynamics	Sort engine efficiency from the smallest to the largest	Check the high reservoir temperature if the low reservoir is known	Design the lowest temperature of a room
	Determine the temperature on the cold reservoir based on the power generated	Assess the efficiency specifications of a machine	

The results of the preparation of HOTS test items all amounted to 40 items including 5 items as an anchor item. The number of items on the aspect of analyzing as many as 16 items, evaluating 16 items, and creating 8 items. The number of items on the aspect of analyzing as many as 16 items, evaluating 16 items, and creating 8 items. Both A and B have the same problem grid, but different context and sequence of questions. Each package of questions was reviewed by physics education experts. The elements reviewed included material, construction, and language. Based on experts' judgments it is stated that the test device met the category valid and feasible to use, with CVI of 1.00.

Grain specifications were obtained through trial activities. Specification of the items in question includes model fit tests, reliability estimates, and estimated difficulty levels. The first is the model fittest, i.e. the matching of grains with the Rasch model, by looking at INFIT MNSQ and Outfit t value. The Quest Output shows 40 items matching the Rasch model, having INFIT MNSQ between 0.77-1.30 and Outfit t $\leq$2.

Estimates of reliability coefficients at the test stages are 0.92 and 0.97 (Table 2), respectively, so the tests are considered to have very high reliability. This reliability is seen from the output of the Quest program which presents the results of test reliability according to CTT, which is an internal consistency index. At least the reliability coefficient of 0.90 can be used as the basis for decisions about individuals (Suryabrata, 2000). Thus, the developed test device qualifies as a HOTS instrument assessment of the thermodynamic material.

Table 2
Estimated Parameters

Parameters	**Try Out**	**Measurement**
INFIT MNSQ	1.00$\pm$0.09	1.02$\pm$0.07
OUTFIT MNSQ	1.00$\pm$0.09	1.02$\pm$0.07
Reliability	0.92	0.97

The results of the grain analysis gave the output estimation of difficulty for 40 items in a trial of -0.88 $\leq$difficulty $\leq$+1.17 and the measurement result gave an -0.81 $\leq$difficulty $\leq$+0.92 (Table 3) estimation. Based on the requirements of difficulty values according to Hambleton and Swaminathan (1985), then the HOTS question developed meets the required criteria as it is between -2.0 $\leq$difficulty $\leq$+2.0.

Table 3
Difficulty Level of Items

Difficulty Level	**Try Out**	**Measurement**
Highest	+1.17	+0.92
Lowest	-0.88	-0.81
Mean	0.00	0.00
Standard deviation	0.64	0.56

The result of HOTS test device test of thermodynamic material yields the TFI and SEM curves presented in Figure 1. Figure 1 shows that assessment tools provide high information to HOTS learners and low measurement error rates when tested on respondents with capabilities ranging from -2.3 to +1.6. This means that the appraisal tool is more appropriately used on the respondents with the ability between -2.3 to +1.6.

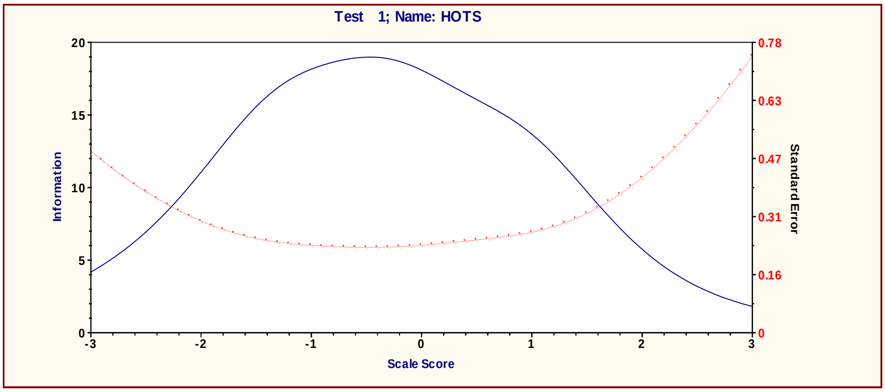

Figure 1: Curve of TFI and SEM

In addition to the total information function in figure 1 above, it can be seen as the minimum ability that the learners must have to be able to correctly answer each item. The rubric rating in this HOTS assessment is 4 if the reason and answer are correct, 3 if the reason is right and the answer is wrong, 2 if the reason is wrong and the answer is correct, and 1 if the reason or answer is wrong. Figure 2 explains characteristics of item number 1, learners can get a score of 1 with minimal ability -3.0, score 2 with minimal ability -1.5, score 3 with minimal ability -0.5, and score 4 with minimal ability +3.0. Each item has different abilities and probability relationships.

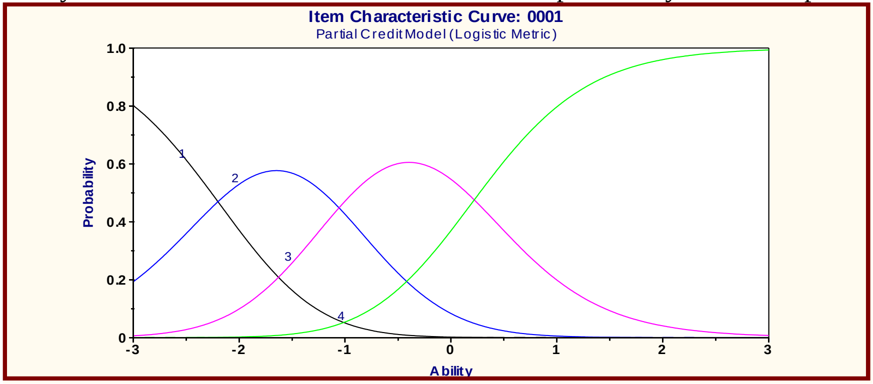

Figure2. Item Characteristic Curve for Number 1

Problems that have been declared eligible to be used to measure HOTS learners include the ability to analyze, evaluate, and create. Thermodynamics material is divided into several sub-material and each of the sub-material is developed according to the HOTS indicator. Each submitter and indicator got a different response from the learner. This is presented in Table 4. Ability to analyze was highest on sub-material 2nd Law of Thermodynamics 50.86%. Ability to evaluate was the highest in the 1st Law of Thermodynamics 41.02%. Ability to create was the highest in 1st Law of Thermodynamics 27.00%. The average learner's ability ranged from the highest to analyze 41.28%, then evaluate 38.70%, and at the lowest is to create 20.02%.

Table 4

HOTS of Students

Sub-material	**HOTS (%)**		
	Analyze	**Evaluate**	**Create**
the Third Law of Thermodynamics	38.55	36.31	25.14
1st Law of Thermodynamics	31.98	41.02	27.00
2nd Law of Thermodynamics	50.86	37.43	11.71
Application of the Laws of Thermodynamics	43.74	40.01	16.25
Mean (%)	41.28	38.70	20.02

The applicable thermodynamic HOTS problem in both content and constructs should be tested for consistency in order for the model to be stable and can be routinely used. According to Sarstedt and Mooi (2014), the reliability of a particular product can be stable when it meets the following qualifications: measurement stability, internal consistency reliability, and inter-rider reliability. Matter of thermodynamic HOTS is to make it fit needs, features novelty and is supported by strong theoretical and empirical ground, to have inter-component consistency (Plomp & Nieveen, 2007).

Differences in the level of ability of learners in their ability to analyze, evaluate, and create caused by several things. First, learners have not been accustomed to responding to questions in the form of multiple-choice options with a closed reason (choice of answers and reasons provided). The reasons presented in the selection contain the appropriate concepts to get answers in solving the problem. Second, HOTS have not been trained and developed optimally in schools. Learners are accustomed to lower order thinking thinking (LOTS), which includes the ability to remember, understand, and apply. As a result HOTS lowers learners, where the higher cognitive domain is given, the results show lower.

The score of HOTS in physics subjects, if associated with the results of national exam scores in Nganjuk district, with an average of 45.11 from a maximum value of 100, and including the low category. But there are some high schools in Nganjuk who are ranked well. This can be related to the differences in high school physics learning strategies in each school. The difference of learning model in school, with more dominant teacher lecture, the more dominant group. However, physics teaching in schools emphasizes how to solve problems in textbooks, not how to develop the concepts of learners so that they can solve physics problems in various cognitive domains.

Conclusion

Based on the description of research results, conclusion are drawn as follows, (a) the content validation results indicate that A and B test devices, each of which contain 20 items with 9 anchor items, through expert judgment have fulfilled the content validation requirements, (b) thermodynamics HOTS test kits have obtained empirical evidence fit with partial credit model, (c) the difficulty level of thermodynamic HOTS is in the range -2.0 to +2.0. The difficulty levels of the most difficult tests in sequence are aspects of analyzing, evaluating, and creating, (d) reliability of HOTS has fulfilled high categorical requirements, and based on the total function of information about HOTS thermodynamics appropriately used to measure HOTS learners with ability -2.3 to +1.6, and (e) the result of the students' response to the ability to analyze is the highest, then followed the ability to evaluate, and create.

References

Adam, R.J. & *Khoo,* S.T. *(1996). Acer Quest: The interactive test analysis system*. Victoria: The Australian Council.

Brookhart, S. M. & Nitko, A.J. (2011). *Assessment and grading in classrooms*. New Jersey: Pearson Education Inc.

Mardapi, D. 2012. *Pengukuran, penilaian, danevaluasipendidikan*. Yogyakarta: MuhaMedika.

Glencoe Science. (n.d.). *Performance assessment in the classroom.* New York: Glencoe McGwaw-Hill.

Hambleton, R.K. & Swaminathan, H. (1985). *Item response theory*. Boston: Kluwer Nijhoff Publishing.

Kozulin, A., Gindis, B., Ageyev, V. S., & Miller, S. M. (2003). *Vygotsky's educational theory in cultural context. Educational Research*. https://doi.org/10.1017/CBO9780511840975

Metcalfe, J. & Dunlosky, J. (2009). *Metacognition.* USA: SAGE Publications Ltd.

Piaget, J. (1976). *The psychology of intelligence [Versielektronik]*.Taylor & Francis eLibrary.

Plomp, T., & Nieveen, N. (2007). *An introduction to educational design research.* Paper presented at the Proceedings of the Seminar Conducted at the East China Normal University [Z]. Shanghai: SLO-Netherlands Institute for Curriculum Development.

Richland, L. E., & Simms, N. (2015). Analogy, higher order thinking, and education. *Wiley Interdisciplinary Reviews: Cognitive Science*. https://doi.org/10.1002/wcs.1336

Saido, G. M., Siraj, S., Nordin, A. B. Bin, & Amedy, O. S. A. (2015). Higher order thinking skills among secondary school students in science learning. *The Malaysian Online Journal of Educational Science*, *3*(3), 13–20.

https://files.eric.ed.gov/fulltext/EJ1085914.pdf

Sarstedt, M., & Mooi, E. (2014). *A Concise guide to market research: The process, data, and methods using IBM SPSS statistics* (2nd ed.). New York: Springer.

....✳ ✳ ✳....

Citation of this Article:

Agustihana, S., & Suparno (2019). Profile higher order thinking skills in Thermodynamics of senior high school students. *Pakistan Journal of Educational Research and Evaluation, 7*(2), 1-10.