



## The Development of Critical Thinking Test Based on Higher-Order Thinking PISA Version in the Historical Learning at Senior High School

Alya Wulan Nur Fatimah\* ; Nunuk Suryani; Sri Yamtinah

Department of Teacher Training and Education, Sebelas Maret University, Indonesia  
Email: [Alya\\_wulannurfatimah@yahoo.co.id](mailto:Alya_wulannurfatimah@yahoo.co.id)

<http://dx.doi.org/10.18415/ijmmu.v5i2.132>

### **Abstract**

This research aims to create critical thinking test as measure tool that valid and reliable to used in the historical learning evaluation at Senior High School. The development method that is use in this research accordance to CIPP evaluation model (Context, Input, Process, Product) that developed by Daniel Stufflebeam and focused to input, process and output. The steps in the development of research method reffer to the development step that has implemented by Aman. I Instrument consists of question device, and the using directive. To get the question with a good category so it is done validation of expert and senior teacher (practitioners). The analysis of score results in the expert validation step uses item that result 25 with a good content construction. Then, it is done the trial test as implementation step. Analyzed of The question properness is analyzed by SPSS for statistic version 17 program. There are 25 questions in analysis result that produce 20 questions that valid and reliable. Therefore, assessment based on higher order thinking PISA version has fulfilled the requirement to measure critical thinking skill of Senior High School Student.

**Keywords:** Assessment; Higher-order thinking; PISA; Critical thinking

### **Introduction**

PISA (Programme for International Student Assessment) is student assessment program with international standard which is held by OECD. PISA study that is held by OECD (Organisation for Economic Co-operation & Development) and Unesco Institute for Statistics to measure student ability in the end age of compulsory education. It is used to know student readiness for face the challenge of knowledge society (knowledge society) nowadays. The assessment that is done in PISA oriented to the future, it is tested the young generation able to use their skill and knowledge in facing the challenge of real life. Not only measure the ability that is written in the school curriculum.

The question of PISA version in the historical lesson is development modal of history question use PISA concept. PISA concept that is used in this research based on literature that will make as student footing constructing learning material.

According to introduction study that is done at SMA N I Sragen states that critical thinking skill of student is still low. It can be seen from such as problem-solving ability in the learning process, written pre-test result and assessment ability (asesmen) that is used. In the learning process, the student is less skilled in solving the problem. It is not caused student incapable but student does not common. Student tendency only gets material from books and a short reading from the internet so they are only learning textbook.

The pretest result that is ditujukan for a student of class XI social science 1 and XI social science 2 at SMA N 1 Sragen support the assumption of the low student critical thinking skill. Questions of level C4 and C5 that should be needed analysis ability that only answered with short answer. The writer assumes that the main problem of low critical thinking skill is placed on the assessment aspect. The assessment that has used still tendency use questions with the low level of C1, C2 and C3.

General paradigm that is formed in the historical subject is boring, history is known with recitation that still make student under estimate with the historical lesson. It is not separated from the assessment that is used. Question level C1, C2 and C3 domination in the historical evaluation activity make assumption of historical subject that identic of recitation is more be entrenched.

Historical learning process experiences progressing from time to time. Start with the model using and learning media that creative and innovative, then entering value from historical moment that aims to create the meaning historical learning. However, it will not be influenced significantly when assessment used still the old model. Therefore, the assessment must be relevant with the use of learning design.

Curriculum 2013 has paradigm of constructivism, that requiring student to develop their own knowledge based on their experiences. Assessment model that uses questions in level C1, C2 and C3 have claimed not relevant yet. The assessment model that is held now, it does not prefer enough because more oriented to learning result. Therefore, the writer needs to develop Assessment model based on Higher Order Thinking (HOT) SIPA version that is assumed relevant between learning and assessment, both of them use constructivism principal.

In the learning, evaluation activity can be created as one step in the remedial effort of education quality. The achieving of education purpose can be by evaluation activity that is done by component of education holding, included teacher. According to the Act No 20 Year 2003, evaluation discuss about National Education System, in Chapter 1 Article 1 verse 21 states that Education evaluation is activity of controlling, collateralizing, and deciding education quality to every education component in every line, level, and education type as the responsibility form of the education holder.

In the learning process, evaluation used as a structural process in order to know the student competency in the learning process based on learning that wants to be achieved. The evaluation of the learning process is done through assessment activity or assessment. The assessment is assessment activity to the measuring result. Assessment activity that is done by the teacher, it will become the step to learning evaluation process generally in a subject.

Based on introduction study that according to the script documents of repeat the end of the semester in the specialization of historical subject of class XI in 2015/2016, it is gained a fact that assessment model is not suitable yet which is expected. From 25 soal multiple choice, it is known the data in Table 1.

**Table 1** Category of soal UAS historical subject specialization on class XI TP 2015/ 2016

No	Amount of question	Question category based bloom theory revision
2, 3, 7, 10, 18, 19	6	C1 (Remembering)
1, 5, 6, 8, 11, 13, 14, 17, 20, 23	10	C2 (Understanding)
4, 9, 12, 21, 24	5	C3 (Implementing)
15, 16, 22, 25	4	C4 (Analyzing)
-	0	C5 (Evaluating)
-	0	C6 (Creating)

The assessment position as a form of learning evaluation becomes very important because it will become the responsibility to the learning process implementation in the curtain time. Therefore, an evaluation must be implemented well with proper procedure, accountability in the learning implementation can show the result that represented the student competency.

Therefore, the connecting or the relevant between each learning components, such as learning process and evaluation activity become very important in order to create the valuable learning. Therefore, it needs to do the research about test development of critical thinking based on higher-order thinking PISA version in the historical learning at Senior High School.

### ***Materials and Methodology***

Development method that is used in this research accordance with CIIP evaluation model (Context, Input, Process, Product) that is developed by Daniel Stufflebeam and focused on input, process, and output. The step in the research method refers to development stage that has practiced by Aman (Aman., 2012), as follows:

First is introduction stage. This stage is the step of collecting information related to the problem reviewed, such as evaluation system of historical learning that ran at Senior High School. Introduction study is done through observation, document analysis, interview with teacher and student, review the works about theories that can be a supporter such as the theory that is related to history and historical learning.

Second is planning stage. This step is planning step of historical learning evaluation model at Senior High School and the instrument planning of data collector as well as assessment model. Besides that, in this step is arranged tested design as the development of product draft/ the beginning shape product. This design method is still hypothetical. It is said hypotetic because the effectiveness has not proven yet and can be known after testings (Sugiyono., 2015).

The third stage are trial, evaluation and revision. The testing should be done to know assessment device model based on HOT, PISA version that is developed more effective compared with model and assessment that has implemented. Furthermore, it is used for guarranteeing the product quality then recommended the using.

The instrument is measure tool that is used to get the information about what will be measured, with another word in order to get measure result. A thing that should be concerned in the measuring process is deciding the proper measure tool. Suitable is meant that can give information and the proper measuring result after the measuring.

According to Siregar (Siregar., 2014), validity that shows a measurement tool can measure what want to be measured (a valid measurement if it successfully measure the phenomenon). Sudaryono (Sudaryono., 2012) explains about the testing technique of test validity in the learning result through two ways, they are: 1) The analysis that is done with the rationally thinking way or the analysis uses logical (logical analysis), and 2) Empirical analysis: the analysis that is done basing yourself to the empirical reality, which the analysis is done with use empirical analysis (hlm.140).

Meanwhile, empirical validity accordance to Sudaryono (Sudaryono., 2012) means as validity that is decided based on criteria, both internal criteria (that test is criteria) and external criteria (test measure result outside that test that becomes the criteria). Learning result test is said that has empirical validity if based on the analysis that is done to the observation data. it is proven that the test properly can measurement learning result that should be measured.

Product validation in this research uses test validity rationally and is done by the searching from two sides that is explained above. Instrument validity goal in the assessment development in this research is done by some of the expert validation and senior teacher as practitioners.

Haynes et. al states that the meaning of content validity is how further the elements in an measure instrument very relevant and the representation from contract that is suitable with the measuring purpose; Ley states that the content validity is how further properness a test as a sample from item domain that will be measured so the position become important, especially for the measuring of achievement (Azwar S., 2013) explains that the content validity concept covered the meaning of face validity and logical validity. Face validity as a part of the content become the starting point in the deciding of test quality that has a meaning of achieving step in general to the test measuring function.

Expert testing and senior teacher are qualitative and judgemental and is done by an expert panel to know how high the dealing between experts who do the properness assessment of an aitem will be estimated and quantified, then the statistic made as an indicator of item content validity and test content validity (Azwar S., 2013).

Validation test in this step uses formula Aiken's V to count content-validity coefficient that considered to the assessment result from the expert panel and people to the item from how further the aitem is represented the measuring contract. In this case, it is represented the measuring construct, it is meant the relevant aitem is relevant with behavior indicator because behavior indicator is an operational translator from the measuring attribute (Azwar S., 2013).

## **Result**

Statistic Aiken's V formulated as follows: this research result can be explained based on development step that used. Azwar (Azwar S., 2013) explains the assessment technique with giving a number between 1 (it is not representative or is not very relevant) to 5 (it is representating or very relevant). Statistic Aiken's V formulated as follows:

$$V = \sum s / [n(c - 1)]$$

Explanation:

$$s = r - l_0$$

$l_0$  = Validity Assessment number that the lowest (in this case = 1)

$c$  = Validity Assessment Number that the highest (in this case= 5)

$r$  = Number that is given by an estimator

In the Aiken's V formula if the range number V that is gained between 0 to 1.00 means that item has a good content validity generally (Azwar S., 2013). Based on the count of expert validation uses Aiken's V formula is gained the result of 25 question stated valid. It is caused range number V that is gained around 0 to 1.00 so it shows that item has good content validity and support test content validity generally.

Operational trial as implementation step of assessment model based on soal higher-order thinking PISA version through CTL model is done to one teacher and 31 students from SMA N 2 Sragen School Year in 2016/2017.

**Table 2** Assessment result of assessment instrument based on HOT, PISA Version

No	The assessment aspects	Score (%)	Conversion	Criteria
1	Sentence arrangement	91.88	A	Very Good
2	Question Meaning	86.88	B	Good
3	The using Istilah	89.38	A	Very Good
4	Command to do soal	91.88	A	Very Good
5	Not mistake of grammar, speeling and punctuation	90.00	A	Very Good
Avarage		90	A	Very Good

Source: Assessment sheet in operational trial by students

Table 2 shows that the whole average score of assessment instrument is 90 % included category "A" or has criteria "Very Good".

**Table 3** The result of assessment instrument based on HOT, PISA Version

No	The assessment indicator	Score (%)	Conversion	Criteria
1	Question is suitable with KD (the basic competency)	100	A	Very Good
2	Question is suitable for indicator	100	A	Very Good
3	Question can measure critical thinking skill	80	B	Good
4	Clear of question meaning	100	A	Very Good
5	Clear of the comand to do questions	100	A	Very Good
6	The clear of term	80	B	Good
7	The good sentence	100	A	Very Good
8	There is no mistake in grammar, spelling and punctuation	80	B	Good
9	Correct Answer Key	100	A	Very Good
10	Objective scoring	100	A	Very Good
11	Enough time for student do the questions	100	A	Very Good
Average		94.55	A	Very Good

Source: Assessment sheet in the operational trial by teacher

Table 3 shows that the whole average score of instrument indicator is 94.55 % and included category “A”. Based on two tables above, the whole avarage score of assessment indicator is 92.275 and included to category “A”. Therefore, assessment of based soal higher-order thinking PISA version in the historical learning fulfill the conditions that will use to measure critical thinking skill of Senior High School Student.

### ***Discussion***

Human Resources that a quality can be described as the competences of individual and has skill, one of the skill is thinking critically. Grieco research (Grieco., 2016), states the importance of improving the ability of student critical thinking. Then the research of Ricketts (Ricketts., 2003), “Critical thinking skills of selected youth leaders: the efficacy of leadership development, critical thinking dispositions, and student academic performance”, state that there is a positive correlation between work (experience or exercises) and the ability of critical thinking. The ability of a good critical thinking, it will be relevant to the high thinking skill.

History is a science that learned about the past moments. History as discipline science consists of moment chain that rich of meanings. Historical learning is not only oriented to the science aspect but

more pressure to the deep meaning in the moment which happened in the past. Therefore, from historical learning hoped that student can understand and take values from what they learn. Understanding process and take values from the learning material, it needs the high ability to think such critical thinking skill.

This research in line with the research that has done before, it is assessment "Relevant Based on Higher Order Thinking PISA version in the historical learning in the improving Critical Thinking Skills of Students at Senior High School". This research result states that the components in the learning design should be had relevant each other. The obstacles that is faced in the learning model implementation of problem-solving has not provided that can measure validly. Therefore, it is important to develop the relevant instrument to measure the achieving level in the learning process. The measure tool is assumed that should mention constructive components, such as the assessment based question of higher order thinking PISA version.

Therefore, the relation and the relevant between each learning components such as the learning process and evaluation activity become very important to create the meaning learning. It is suitable with the research result that science or learning meaning only can be achieved when the student can develop their knowledge based on the experiences that suitable with constructivism principles (Baron., 2016). Then the research (Legare., 2002) strengthen that constructivism approach gives more opportunity. It involves critical thinking skill that the method that centered to teacher because they usually offer new problem or the confusing situation.

This research result shows that assessment development based on HOT answer the problem. The problem is the limiting of assessment devices that can be used by practicion in the education world, the especially historical teacher at Senior High School level in Sragen.

Validity analysis of soal use *iteman* with formula of Aiken 's V shows that the whole soal has good category shown with the counting acquisition V range 0.00 s/d 1.00. It is meant that assessment empirically based on higher-order thinking PISA version, it can be stated valid. Test that contains a high level of thinking questions that valid, it is encouraging students to think deeper. Beside that, material construction that is arranged with adopting literation concept on PISA questions, it can push students to think critically with connecting the reading material with harmonizing on the problem that must be finished.

Furthermore, assessment reliability tested based question HOT, PISA version uses the help of SPSS application for statistic version 17 shows that instrument has the high reliability, that is 0.736. The output of count result can be seen in Table 4.

Assessment reliability based soal higher-order thinking PISA version includes the high category because it placed on the range 0.600-0.799. This is confirmed that measuring result uses this instrument is reliable. According to reliability coefficient, test information function, and estimation parameter meant assessment based question of higher-order thinking PISA version is reliable and has the high stability.

**Table 4** Case processing summary

	N	%
Valid	30	100,0
Excluded <sup>a</sup>	0	,0
Total	30	100,0

Reliability statistics

Cronbach's Alpha	N of Items
,736	26

### **Conclusions**

Learning components that consist of some aspects must have relevant to make the meaning learning. Constructivism Paradigm become education development basic is the paradigm that emphasizes process than a result, and require the student to develop their own ability by relevant with their experiences.

Assessment model based on HOTS historical PISA version emphasize student in the thinking process when finishing a problem. Model design is arranged as an effort to train student the mindset in order to have mindset that analytic critical.

From the explanation above, assessment model based on HOTS historical PISA version is claimed suitable if it is implemented in the historical learning evaluation. Learning evaluation is a system that is arranged systematically that is used to know the competency achieving in the curtain time. Evaluation process must be done correctly related to the principals and indicators that want to be achieved. It is intended in order of the evaluation process can represent the competency that had by students.

Even though, it has not optimal yet and there are still some notes that need to be repaired in the next research, but this research has result from critical thinking test as assessment devices based on of high-order thinking PISA version that can be utilized by education practices an especially historical teacher at Senior High School.

### **Acknowledgement**

Thanks to the Ministry of Education and Cultural Indonesia that has given scholarship “Masyarakat Berprestasi” to us as a student in Historical Education Study Program of postgraduate at University of Sebelas Maret Surakarta. We feel very helped with this scholarship.



## **References**

- Aman. (2012). Pengembangan Model Evaluasi Program Pembelajaran Sejarah Di SMA. *Jurnal Penelitian dan Evaluasi Pendidikan*, 437-456.
- Azwar, S. (2013). *Reliabilitas dan Validitas*. Yogyakarta: Pustaka Pelajar.
- Baron, P. (2016). A Cybernetic Approach to Contextual Teaching and Learning. *Constructivist Foundation*. p91-100.10p.
- Grieco, L. N. (2016). Meeting the Demands of the 21st Century Workplace: Effects of Critical Thinking Instruction on the Application of Critical Thought.
- Legare, G. (2002). An investigation of the effect of task design on the development of critical thinking skills by engineering students.
- Ricketts, J. C. (2003). Critical Thinking Skills of Selected Youth Leaders: the Efficacy of Leadership Development, Critical Thinking Disposition, and Student Academic Performance.
- Siregar, S. (2014). *Statistik Parametrik untuk Penelitian Kuantitatif: Dilengkapi dengan Perhitungan Manual dan Aplikasi SPSS Versi 17*. Jakarta: Bumi Aksara.
- Sudaryono. (2012). *Dasar-dasar Evaluasi Pembelajaran*. Yogyakarta: Graha Ilmu.
- Sugiyono. (2015). *Cara Mudah Menyusun Skripsi, Tesis, dan Disertasi*. Bandung: ALFABETA.

## **Copyrights**

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).