

Developing a Critical Thinking Test in Social Science for Grade V Elementary School

Kurniawati^{1*}, Edi Istiyono²

¹Department of Primary School Teacher Education, Universitas Negeri Yogyakarta, Indonesia

²Department of Educational Research and Evaluation, Universitas Negeri Yogyakarta, Indonesia

*Corresponding author. Email: kurniawati@uny.ac.id

ABSTRACT

The purpose of this research is to describe the quality of a developed critical thinking assessment instrument in Social Sciences for Grade V elementary school. The research population was the grade V students of state elementary schools which have already implemented the 2013 Curriculum in Yogyakarta city with a sample size of 218 students. The analysis was done by using the item response theory approach through BILOG-MG 3.0 software program. The critical thinking test instrument in social sciences for grade V consists of 30 multiple-choice questions. The critical thinking test instrument has 30 items fitting with the 2 PL model ($p > 0.05$). The average difficulty level of the items is -0.71 while the average of the discrimination index is 0.52. This test provides information on the maximum of 6.09 and its smallest error of measurement is 0.40.

Keywords: *assessment, critical thinking, elementary school, social sciences, test development*

1. INTRODUCTION

Currently, there are various levels of competencies in different fields that carry out the impact of the qualified human resources requirement. To reach the qualified human resources, a qualified education is needed. One of the references to improve the education quality can be seen from the qualified output. A good education can be hoped to develop some important skills in life. The demand for having the skill in 21st century becomes the basic need for human beings.

Some 21st-century skills [1] are: adaptation, complex communication skills, problem-solving skills, self-managing development, and thinking system [2]. There are also other skills such as [3] spoken and written communication, critical thinking and problem solving, professionalism and ethics, various teamwork, applying technology, leadership and project management.

To settle the challenges of the 21st century, the Indonesian government made some changes to the curriculum. The 2013 Curriculum aims to develop productive, creative, innovative, and effective human resources that can be useful and give contribution for the mass. It is firmly stated [4] the objective of the 2013 Curriculum is to elaborate the students' ability in observing, asking, logic, and communicating the topic discussion to others.

A 21st century skill that is important to build since childhood is critical thinking. Critical thinking can determine the success of a person's life. It connects to how people can solve their problem. 21st-century skills are not the demand or responsibility for one subject matter at school. Critical thinking skill is not only important for science, but also for others knowledge and to solve everyday problems. Critical thinking skill is not only needed for intellectual development but also give value of life in nature and social life. Critical thinking skill will grant the students to be a good problem solver in nature and social life. As important as this critical thinking skill in science or in social learning so this skill is needed to be elaborated since childhood. Since the children are in elementary school to university. One of the subject matter that can be hoped to be able to develop the critical thinking skill is social science. Social science has the responsibility because it directly connects to social life which is done by the students.

In addition to the learning process, the assessment and the form of assessment that is used by the teachers will influence the critical thinking of the students. The assessment can be used to help students increase their thinking skills. It is in line stated that assessment can be applied in the class to increase their critical thinking [5]. The questions in high-level thinking are useful for

the students to improve the material. So it can be concluded that a test with high critical thinking will stimulate the students to improve that skill [6].

Critical thinking skill has not been expanded at schools. Schools still stress memorizing where the students are given one right answer [7]. The learning process does not give the chance to the students to expand their critical thinking like creating new ideas, giving opinions etc.

Students (grades 4 and 5) sign negatively to social science than mathematics and science. They have an image that social science is a difficult lesson, boring, unworthy, and irrelevant in real life [8]. Many students have negative behavior towards social science and they do not understand the importance of social study or its connection in their life [9].

Setiadi [10] found the implementation of the 2013 curriculum with assessing students by the teachers. Teachers usually ignore to make the blueprint, it makes the items uncontrollable so they cannot measure to reach the objective. The teachers also do not analyze the instrument before doing the assessment, whereas it is important to do this because it has the role to get a valid and reliable instrument.

Based on the other researches on critical thinking skills and its assessment that is used to show the social science assessing process in Yogyakarta tends to measure low level of thinking which does not give the assessment to measure the development of students' critical thinking. Based on the interview result with grade 5 teachers in state elementary schools Yogyakarta in arranging the items, they never develop the items with high order thinking items, especially to measure the critical thinking skill.

Besides the study of the portion item used, this study also discussed about the items in multiple choice test form with the content by identifying the use of figure, table, and chart. The identification result shows that it is only a few items that used the stimulation. The use of figures, tables, and charts is one of the methods to create a multiple-choice item that can measure students' critical thinking skill [11]. The choice of the same type of item every year causes the students to memorize it. It makes the items only measure the reminding aspect. This multiple-choice item will cause students to memorize the concept and formulate what is in their book.

Based on several researches, it is known several factors are needed: 1) the assessing instrument has to cover the learning process in The 2013 Curriculum, that is developing students' critical thinking skill; 2) the assessing instrument that is easy to be used by the teachers and also students; 3) the valid and reliable assessing instrument or fulfill the standard measurement. From those three needs, so it is necessary to develop a multiple-choice test that can measure students' critical thinking skill. This multiple choice test development will enhance the variation of item and it can be used in assessing and accommodating the critical thinking skill.

2. LITERATURE REVIEW

2.1 Critical Thinking Skill In Social Studies

Critical thinking skill is important in every field of knowledge to solve the problem they faced every day [12]. Critical thinking skill is not only needed for intellectual development but also give the value for daily life. Accordingly, critical thinking skill will grant the students to be a problem solver and can apply it in various phenomenons.

Critical thinking skill is important to be developed to the students, it is because if they were adult they would act actively in mass. They can solve the problem in mass, for example they can give an idea for the happening problem. In order to improve the students' adaptation with the environment so this critical thinking skill can be elaborated through social science study.

Critical thinking skill as a future competency gives the consequence that is the social science learning process must pay attention to critical thinking skill. The other important aspect is assessment. Teachers must develop the critical thinking skill and also be supported by the good assessment. Teachers must be able to set a continuity assessment that can observe the improvement of students' critical thinking skill, starting from the beginning of a lesson, process, and ending the lesson. Students' assessment is useful as a feedback where it can be based on the work and product.

2.2 Social Studies

In this globalization era, social science study is not only settled to give the knowledge for students. Nowadays, the focus is how to transform social science from the memorizing skill to the developing logic skill for the students. Logical learning is developed in order to increase their reflective thinking. Increasing reflective thinking becomes the most important thing in a learning process.

According to Ellis, social science according to NCSS has the main objective to help the students so that they can improve their skill in making decision and rational for the public importance as a good citizen, the various democratic citizen and interdependent to each other [13]. By having a critical thinking skill, students will be easier in inquiring and make them be able to build their own knowledge.

3. METHOD

3.1 Data Collection

This research is a development research study employing a quantitative approach, which aimed to create a product. The product in this research is the instrument of social science critical thinking test for fifth graders in the state elementary school. This product is created by some procedures of instrument development. The instrument development that is applied in this research used the theory of developing a

test that consists of nine steps, those are (1) constructing the specification of a test, 2) writing the question, 3) analyzing the question, 4) executing the tryout, 5) analyzing the items, 6) fixing test, 7) compiling test, 8) doing test, and 9) interpreting the test. This instrument of developing test is based on the need analysis and theoretical study of social science critical thinking [14].

This research was conducted on January 2017 until September 2017. The beginning of developing test that is test preparation, validation, and test compilation is conducted on January 2017. This research was accomplished in the state elementary school in Yogyakarta. The research subjects were all the fifth grade students in Yogyakarta that has applied The 2013 Curriculum. This study has done the try out at 218 students before doing the research.

3.2 Data Analysis

The data were analyzed using *Item Response Theory* and were computed by BILOG-MG program. The result of this research is a document of the analyzing items that will give the information of the parameter estimation that is item difficulty and item difference. The result will also give the information of the students' critical thinking in the social science study.

4. RESULT

4.1 The Result of Test Development

The first product of this study is a set of social science critical thinking test for students grade 5 of the state elementary school Yogyakarta in semester two in the academic year of 2016/2017. The specification of the items is based on the core competency and the basic competency and then it is developed into some indicators in the blueprint of the test. The materials tested are theme 6, theme 7, theme 8, and theme 9, where this study is only focused on the basic competency for social science study. Furthermore, the items test which has developed is discussed by some experts in the field of measurement, social science study in elementary school, and teacher to get the expert judgment until this instrument is decent to be employed.

Table 1. Item distribution of social science critical thinking test for fifth graders

Aspect	Theme 6	Theme 7	Theme 8	Theme 9
Interpreting	1	2, 3, 4	5, 6	8
Analyzing	9, 10, 11	12, 13, 14, 15	7	17, 18, 19, 20, 21
Evaluating	-	22, 23	24, 25	26
Concluding	-	27, 28	29, 30	-

The material of this instrument test of social science critical thinking consists of themes 6, 7, 8, and 9. In every theme that is developed contains four indicators

of critical thinking that covers the aspect of interpreting, analyzing, evaluating, and concluding. The use of this critical thinking aspect is in line with Rosnawati [15]. The items of this critical thinking test are expanded in the form of multiple choices with four alternative answers.

4.2 The Try Out Result

Before executing the items analysis using IRT, the earlier step is doing *unidimension* assumption test, local independency, and also parameter invariance [16]. *Unidimension* assumption test proved that every single item can only measure one skill. Based on the factor analyzing testing, the result is presented in Figure 1.

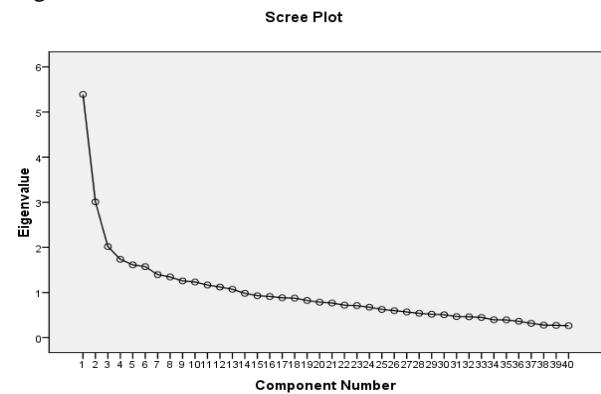


Figure 1 Scree plot of the result of exploratory factor analysis

Based on the scree plot in figure 1, it shows that the range of component 1 to component 2 is far, while the range of component 2 to component 3 and so on is near. Beside of that, it can be seen in figure 6 that shows the eigen value is starting to slope at component 3. It shows that there is 1 dominant factor than others that gives much more contribution to the variance which can be explained. A test can be stated as a unidimension test while component 1 and 2 in the scree plot (Figure 1) has an adequate [17]. Based on the scree plot in Figure 1, it could confirm the conclusion that is the instrument of critical thinking test for social science tends to unidimension.

The assumption of the local independency is automatically fulfilled while the assumption of unidimension has fulfilled [16]. Based on the theory, this instrument has fulfilled the assumption of local independency. The third assumption that has to be fulfilled is invariance of the parameter. Invariance of parameter can be tested by dividing 218 samples into two groups, the odd group and the even group. The result of the items parameter estimation from each sample can be seen in the scree plot and after that it could be correlated. If the correlation is positive and high so the invariance parameter assumption of the items is fulfilled [18]. The result of invariance parameter assumption of the items can be seen in Figure 2.

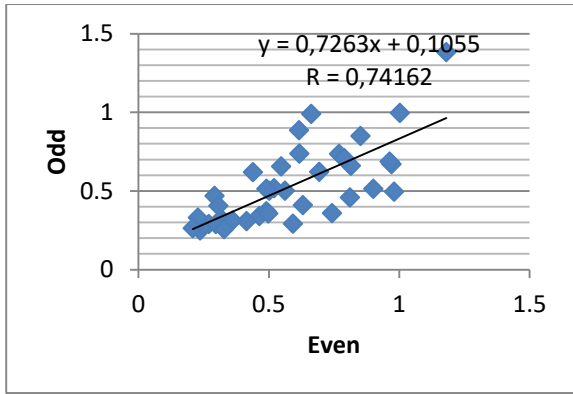


Figure 2 Scree plot of the invariance parameter assumption in item difference test

Figure 2 described the scree plot of the invariance parameter assumption in item difference after doing the test in the odd items and even items. In Figure 2, there can be seen the estimation scores is spread and relatively located near the linear line. The correlation between item difference in odd items and even items is high enough that is 0.74162. Based on the scree plot and the correlation, it can be concluded that the parameter of item difference is invariance. Therefore, the parameter of item difficulty analysis is also conducted by separating two groups into odd and even group by using *Microsoft Excel* so that there can be gained the correlation in Figure 3.

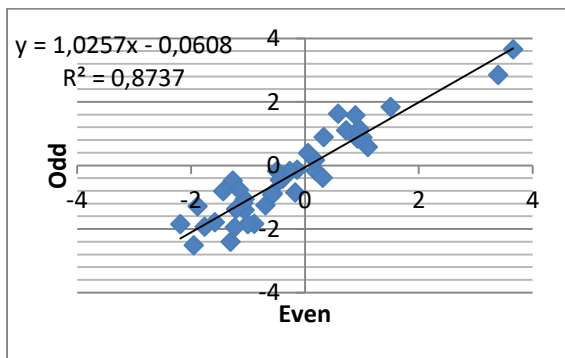


Figure 3 Scree plot of the invariance parameter assumption in item difficulty

Figure 3 shows the scree plot of the invariance parameter assumption in item difficulty after doing the test in the odd and even items. In Figure 3, the estimation scores are spread and relatively located near the linear line. The correlation between item difficulty in odd items and even items is in high category, 0.8737. Based on Figures 2 and 3, the invariance parameter estimation for item difference and item difficulty can be fulfilled as well.

Thus, the ability of the invariance parameter can be tested by dividing the group into odd and even groups. Then, the result of the estimation should be created in the scree plot and make the correlation. If the correlation is positive and high, the assumption is fulfilled [18]. The whole estimation of students' ability in doing the test is in the form of scatter as in Figure 4.

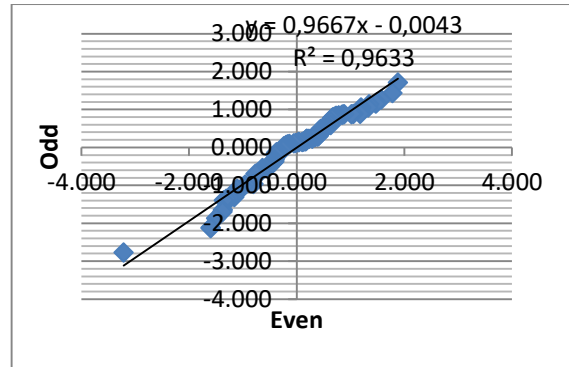


Figure 4 Scree plot of the ability invariance parameter estimation

Figure 4 describes that the invariance of students' ability estimation after it was analyzed by comparing the students' ability in odd and even groups. The result of the scree plot can also explain about the estimation score located near the linear line with the correlation is 0.9633 and it is a very good category. Because of that, there can be concluded that the assumption of the ability invariance parameter is also fulfilled.

The third assumption for IRT analyzing has been fulfilled as well, so it can continue to the next fit model testing to analyze the test that has developed. The fit model testing for 1-PL, 2-PL, or 3-PL is conducted by comparing X^2 scores. The probability score in each item has to fulfill $p > 0.05$, if there is not, so it needs to make a revise before doing the instrument testing.

Table 2. Fit Model if critical thinking test

	1PL	2PL	3PL
Items that fits model	1,3,4, 7,8,13, 14,15,16, 17,18,19, 20,21,22, 23,24,25, 28,29,30, 35,37,38, 39,40	1,2,3, 5,7,8, 9,10,11, 12,15,16, 17,18,19, 23,24,25, 26,27,28, 30,31,32, 33,35,37, 38,39,40	1,2,3, 5,7,8, 9,10,11, 12,16,17, 18,19,23, 24,25,26, 27,28,30, 31,32,33, 35,37,38, 39,40
Fit	26	30	29
Unfit	14	10	11

Based on the result in Table 1, it can be confirmed that the model which is the fittest for this instrument is in 2-PL model. The fitness of 2-PL model for this critical thinking test which has developed because the number of the item test is more fitter to the 2-PL than 1-PL or 3-PL. The model that fits to be applied in analyzing is 2-PL, so the parameter that needs to be noticed is item difference (a) and item difficulty (b) for each item. The items which do not fulfill the good criteria based on those parameters will be discarded from the final product.

The parameter is used in 2-PL model is item difficulty (bi) and item difference (ai), while guessing (ci) is supposed to be zero. Therefore, the fit items with 2-PL model will be joining in the next analysis. Those

items are 1, 2, 3, 5, 7, 8, 9, 10, 11, 12, 15, 16, 17, 18, 19, 23, 24, 25, 26, 27, 28, 30, 31, 32, 33, 35, 37, 38, 39, and 40. The other items that is not fit with 2-PL will not be joining to the next analysis although they may have the good item difficulty and item difference. Those items are 4,6,13,14,20,21,22,29,34, and 36. The result of the items analysis by using 2-PL, there can give the information about the characteristics of item difference (a) and item difficulty (b). The result is shown in Table 2.

Table 2. The result of the analysis of items characteristics

Items	a_i	b_i	Items	a_i	b_i
1	0,476	-0,861	16	0,467	-1,11
2	0,19	-0,347	17	0,547	-1,852
3	0,478	-0,299	18	0,512	-0,78
4	0,531	1,039	19	1,121	-1,017
5	0,486	-1,135	20	0,624	-1,849
6	0,558	-1,229	21	0,558	-0,99
7	0,755	-1,451	22	0,701	-0,803
8	0,297	-1,981	23	0,6	-0,924
9	0,318	-0,662	24	0,62	-0,791
10	0,666	-1,391	25	0,463	-1,458
11	0,487	-0,611	26	0,608	0,333
12	0,41	0,653	27	0,462	-0,785
13	0,356	0,78	28	0,339	0,623
14	0,547	-0,845	29	0,522	-0,313
15	0,708	-0,597	30	0,291	-0,714

Based on the result of item analysis in Table 2, it can be gained information about the characteristics of item difference (a_i) and item difficulty (b_i). The result shows that all items from number 1 to 30 have the index of item difference (a_i) that belongs to a good category, it can be seen from the score of all the items belonging to good criteria, that is 0 until 2. The lowest index of item difference is in item number 3, that is 0.19 and the highest index of item difference is in item number 20, that is 1.121. It is known that there are 28 items have the index of item difference more than 0.3, that means the 28 items can make a distinction of the low ability and high ability in the condition where the ability is the same as the item difficulty.

The item difficulty for this social science critical thinking test considers in good category. All of the items has the index of item difficulty between -2 to 2. In this instrument, the lowest index of item difficulty is in item 8, that is -1.981 and the highest index of item difficulty is in item number 4, which is 1.039.

4.3 Discussion

The instrument that has developed is in the form of multiple choice because the students have habituated by that form. The sdvantages of multiple choice test is that it can measure the level of thinking from low to

high level, has a large scope of material, and has an objective scoring because it can use computer [14]. Because of that, critical thinking test can also be developed in the multiple choice form.

The blueprint was arranged based on the aspect and sub-aspect of critical thinking skill, thus it was used to set the items. This instrument consists of a set of test with 30 items. After that, the instrument has to be validated by the expert of measurement, social science study, and teachers. The result of validation was computed using Index Aiken’s V, and there are 40 items that will be computed. All of the items are valid, the value is >0.75, 7 items have the index Aiken of 0.75, and the other 33 items have>0.80. it means that the coefficient is high enough or the content validity is good and supports the content of the test wholly [19]. the result of Index V Aiken’s can give the description about the instrument of critical thinking test has been valid and it can be used to measure the ability of critical thinking skill of fifth grade students.

The reliability test estimation of social science critical thinking is high (0.8237). It means that this instrument is reliable in measuring the ability of social science critical thinking for fifth grade students in Yogyakarta. The reliability estimation is 0.82, whichcan be used to measure the students in the class [19] the informational test function is in average category that fits to the ability -4 to 2.75. This instrument is reliable enough because it is arranged from the items which has high informational function [20]. Based on the reliability coefficient, the informational function and the parameter estimation so the social science critical thinking test can be declared that is reliable and has enough stability.

5. CONCLUSION

The conclusions for this research are 1) the instrument of social science critical thinking test for fifth grade which is developed by the researcher consist of 30 multiple choice items by concerning on the indicator of critical thinking in the ability of interpreting, analyzing, evaluating, and concluding 2) the instrument of social science critical thinking test for fifth grade has fulfilled the estimation of reliability, the reliability coefficient is 0.8237 and it is consider in high category, it shows that this instrument is reliable to measure the students’ critical thinking skill in the state elementary school of Yogyakarta. 3) the instrument of social science critical thinking test for fifth grade has 30 items fit by using 2 PL Model, 4) all of the items in this instrument is in the good category by the average of item difficulty is -0.71223, 5) all of the items in this instrument is in the good category by the average of item difference is 0.523267.) based on the information function, the instrument of social science critical thinking test for fifth grade fits to be applied to measure the ability (θ) that is -4 to 2.75. Therefore, this instrument is appropriate to be applied to measure the students’ critical thinking skill in social science that has average and low ability

ACKNOWLEDGMENT

We would like to express our deepest gratitude to everyone who has provided various assistance for the implementation of this research.

REFERENCES

- [1] T. Wagner, *The global achievement GAP*. New York: Basic Books, 2008.
- [2] National Research Council. *Assessing 21st Century Skills: Summary of a Workshop*. J. Washington, DC: The National Academies Press, 2011.
- [3] Trilling & Fadel. *21st-century skills: learning for life in our times*. United States of America: Jossey-Bass, 2009.
- [4] Mulyasa, H, E. *Pengembangandan implementasi the 2013 Curriculum*. Bandung: Rosdakarya, 2013.
- [5] E, Istiyono, M. Djemari, & Suparno. *Pengembangan tes kemampuan berpikir tingkat tinggi fisika (PysTHOTS) peserta didik sma*. *Jurnal Penelitian dan Evaluasi Pendidikan*, 18, 1, 2014, pp. 1-12.
- [6] J, E. Barnett, & A. L. Francis, *Using Higher Order Thinking Questions to Foster Critical Thinking: A Classroom Study*. *Educational Psychology*, Vol. 32(2), 2012, pp. 201. Retrieved: <https://eric.ed.gov/?id=EJ959633>.
- [7] J, W. Santrock, *Life-span development*. Jakarta: Erlangga, 2012.
- [8] Dundar, S & Rapoport, A. *Elementary students' attitudes toward social studies, math, and science: an analysis with the emphasis on social studies*. *The Councilor: A Journal of the Social Studies*. Vol 75, 2014.
- [9] Y. Zhao, & J.D. Hoge, *What elementary students and teachers say about social studies*. *The Social Studies*, Vol. 96(5): 2005, pp. 216-222.
- [10] Setiadi. *Pelaksanaan penilaian pada the 2013 Curriculum*. *Jurnal Penelitian dan Evaluasi Pendidikan*. Vol. 20(2), 2016.
- [11] A. J. Nitko, & Brookhart. *Educational assessment of student*. Boston, MA: Pearson/Allyn & Bacon, 2011.
- [12] Moore, B. & Standley. *Critical thinking and formative assessment increasing the rigor in your classroom*. New York: Eye on Education, Inc. 2010.
- [13] A, K. Ellis, *Teaching and learning elementary social studies*, (6th Ed). Boston: Allyn and Bacon, 1998.
- [14] D. Mardapi, *Teknik penyusunan instrumen tes dan nontes*. Yogyakarta: Mitra Cendekia Offset, 2008.
- [15] Rosnawati, *A formative assessment model of critical thinking In mathematics learning in junior high school*. *Research and Evaluation in Education Journal*. Vol. 1(2), 2015.
- [16] R. K. Hambleton, H. Swaminathan, & H. J. Rogers, *Fundamental of item response theory*. Newbury Park, CA: Sage Publication Inc. 1991.
- [17] R. M. Furr, & V. R. Bacharach, *Psychometrics an introduction*. Thousand Oaks: Sage Publications, Ins. 2008.
- [18] H. Retnawati, *Teori respons butir dan penerapannya*. Yogyakarta: Nuamedika, 2014.
- [19] S. Azwar, *Reliabilitas dan validitas*. Yogyakarta: Pustaka Pelajar, 2015.
- [20] R. K, Hambleton, & H. Swaminantha, *Item response theory*. Boston, MA: Kluwer Nijhoff Publishing, 1985.