Developing Critical Thinking Skills Test for Undergraduate Biology Student

Widi Cahya Adi[1], Hadi Suwono[2], Endang Suarsini[3]
Universitas Negeri Malang, Indonesia
widicahyaadi@gmail.com[1], hadi.suwono.fmipa@um.ac.id[2], endang.suarsini.fmipa@um.ac.id[3]

Abstract: Critical thinking is the main capability that must be owned by the people in 21stcentury. The importance of critical thinking skills is must be prepared for the students as young age,including the Biology students as candidate researcher.Critical thinking skills are important for a researcher to be able to solve problem and make right decision. This research aims to developing instruments to measure critical thinking skillsbiology students. The Instrument developed refer to indicators of Watson-Glaser Critical Appraisal Thingking (WGCTA). Stages of development by make the test based of critical thinking indicators, trial test, and analysis of the test. The results of trial test with 90 students biology concluded that instrument critical thinking skills which consists of 23 multiple choice question this is valid and reliabel, so that it can be used to measure the critical thinking skills of biology students.

Keywords: critical thinking, instrument, WGCTA, biology student

The 21st century is an era of globalization and internationalization (Osman et al., 2010; Fong et al., 2014), people will face competition and global issues, so that every person is important to have 21st century skills (Sahin, 2009; Pheeraphan, 2013). Some of the 21st century skills are critical thinking, creativity, communication, collaboration, communications technology mastery, and life and career skills (The Partnership for the 21st Century Skills, 2007).

The main skill that must be owned by people in the 21st century is critical thinking (Jerald, 2009; Thompson, 2011; Fajrianthy et al., 2016), because it associated with the problem solving process (Friedel 2008) that occurred in the daily life, work, and all other aspects of life (Slameto, 2014). The importance of critical thinking skills can help people to succeed in their life and work (The Partnership for the 21st Century Skills, 2009; Moses et al., 2012; Turiman, 2012; Dass, 2014; Ay et al., 2015) so, the critical thinking skills need to be prepared especially for the young generation (NCREL and Meitri Group, 2003; The Partnership for the 21 st century skills, 2009; Rodzalan and Time, 2015) including the Biology students.

Biology nowadays has great potential in contributing to resolve global issues such as health, food, energy, and environment (AAAS, 2011; Osman et al., 2012). Therefore, it is important Biology students as candidates of biological researchers to develop critical thinking skills so that students can solve problems in biological life. The importance of critical thinking skills for Biology students can shape the process of thought into fast, accurate, and free of assumptions, especially when they are able to solve problems critically and to determine decisive decisions and appropriate actions (Caroselli, 2009), critical thinking skills are also the basic skills in problem-solving in the research (Thompson, 2011; Facione, 2011).

Nowadays, the importance of the development in critical thinking skills for the Biology students have several obstacles, that critical thinking skills are rarely taught officially in class (Addy et al., 2012). In addition, according to the observations of researchers in measuring skills, critical thinking especially for the Biology students is still rarely implemented because of the absence of a standardized test instruments. According to Fajrianthy et al (2016), the

measurements of critical thinking skills in Indonesia encountered some problems, for example: the context of measurementis quite diverse, the measurements that are likely adapt some tests developed by western countries without testing it first (if this test contains a cultural bias or not), the measurement of critical thinking development in Indonesia is mostly implemented in the educational setting of mathematics and physics.

The measurement problems of critical thinking skills in general has become a controversy of experts in the fields of psychology, philosophy, and education (Fajrianthy et al., 2016). The controversy was caused by definitions and indicators for measuring critical thinking skillswhich are still very diverse (Wagner, 2002). Halpern (1999) stated that even when the definition of critical thinking skillsaccording to experts was very diverse, but basically it has the similar basic principles.

There are some experts who have different opinions about the definition and indicators of critical thinking skills. Ennis (1985) defined that critical thinking skills are as reflective and reasonable thinking which focuses on deciding what to believe or do. It isalso a skill which can be measured by indicators that covers basic clarification (which focuses on a question, analyzes arguments, asks and answers clarification and / or challenges questions), bases for a decision (judge the credibility of a source, Observe, and judge observation reports), inference (deduce and judge deduction, induce and judge induction of make material inferences, make and judge value judgments important factors), advanced clarification (define terms and definitions judge, attribute unstated Assumptions), supposition and integration (suppositional thinking, integrate the dispositions and other skills in making and defending a decision). Facione (2000) then stated that critical thinking skillsare defined as a skill to do judging in a reflective way on what to do or what to believe, which can be measured by indicators that cover analysis, inference, evaluation, deductive reasoning, and inductive reasoning. Furthermore, Watson and Glaser (2012) have also defined the critical thinking skills as the ability to identify and analyze problems as well as it seeks and evaluates relevant information in order to reach an appropriate conclusion, which also can be measured by indicators that cover inference, recognition assumption, deduction, interpretation, evaluation of arguments.

The different understanding of some experts about the critical thinking skills puts many suppression to the description of the indicators than to the fundamental difference. The experts then have agreed that the critical thinking skills basically consist of skills to analyze an argument, to make an either inductive or deductive conclusion, to evaluate and make decisions or solve problems (Lai, 2011). The existence of the controversy makes an important challenge for the university to develop an accurate measurement instrument in reflecting the teaching and learning as well as the practices that have been implemented on the campus. (Stassen et al., 2011).

The definitions and indicators to measure critical thinking skills that are widely accepted and often used are the Watson-Glaser Critical Thinking Appraisal (WGCTA) (Wagner, 2002). Some studies have also shown that WGCTA can be used as a tool for critical thinking skills (Husband, 2006; Ejiogu et al., 2006). WGCTA is a psychometric test of critical thinking and reasoning, these tests measure skills related to the problem solving and decision making in different types of questions (Watson and Glaser, 2012).

Based on the problems described above, it can be concluded that the importance of critical thinking skills for Biology students led to the need of critical thinking skills test instruments for Biology students. The purpose of this research is to develop critical thinking skills test instruments for Biology students to be valid and reliable so that it can produce accurate measuring result.

METHOD

This research is a developmental research that will develop critical thinking skills test instruments for Biology students. Stages of development refer to the stages by Hambleton and Jones (1993) with modifications, as follows.

Preparation of Test Specification

The first phase was done by determining indicators of critical thinking skills that are used as the basis of measurement, formulate the indicator in the form of test items. Based on the results of the review, researchers referred indicators of critical thinking skills by Watson and Glaser (2012), which consisted of five indicators (Table 1.1), as follows.

Table 1.1 Indicators of Critical thinking skills

Indicators	Explanation
Making inference	Assessing whether the inference is "definitely true," "may be true," "may be wrong," or "definitely wrong," according to a statement, or is "not enough data" to draw the inference.
Identifying assumption	Assessing whether an assumption is appropriate with the statement
Deductive reasoning	Assessing whether the deductive conclusion of a statement is true or not
Interpreting argument	Assessing about the conclusion "no doubt" of another statement.
Evaluating argument	Assessing whether an argument is classified as "strong" or "weak."

(Source: Watson dan Glaser, 2012).

This stage also determined the specifications of the context and the form of the developed instruments. The developed test instruments were in the form of test items which wereadapted to the context of Biology in general. Furthermore, the test items were developed in the form of multiple choice tests according to the pattern of WGCTA.

Preparing the Test Items

This phase was done by developing test items from each indicator of critical thinking skills. Researchers developed 10 items for each indicator. Finally, researchers produced a test instrument with a total of 50 items.

Small scale Try-out and Test Items Analysis

This phase was done by trying out the test instruments that were developed as intended to determine the weaknesses in such instruments. The try-outwere administered to the 25 Biology students at State University of Malang. The results of further try-outwere analyzed based on the level of validity, reliability, standard deviation, and the level of difficulty in the test items. The results of the analysis were used as a basis of test improvement.

Test Items Revision

This phase was done by identifying the questions with a low level of validity (invalid).

The revisions were carried out by revising the test items based on the language and the clarity of items. The revision aimed to improve the test items so that it was easy to understand and did not give an ambiguous statement. Thus, when the test items will be tried out again on a larger scale, it will make the results of test items analysis better.

Large Scale Try out and Final Analysis of Test Items

This phase was implemented by trying out the revised test items to the students in a larger scale. The try out was conducted to the 90 Biology students at State University of Malang. A large trial results were then analyzed based on the level of validity, reliability, standard deviation, and the level of difficulty of the test items.

Printing and Distributing Test

This stage was the last stage, which was completely done by printing and distributing tests that had been declared as valid and reliable test.

Findings and Discussion

The test instruments of critical thinking skills for Biology students were developed with multiple-choice forms that refers WGCTA pattern. The context's content used in the test instruments was Biology in general. The test instruments consisted of 50 items, divided into five indicators of critical thinking skills. The examples of the test items for each indicator are as follows.

Indicator 1: Inference
Statement
A greenhouse worker discovered that the chrysanthemum plant located on the edge of the shelf often produces flowers which are shorter than flowers of the chrysanthemum plant located in the center of the shelf. The differences condition of the location between on the edge of the shelf and on the center are the light intensity and the airflow. The light intensity and the airflow on the center of the shelf are lower than on the edge of shelf.
Assumption

The chrysanthemum plant located in the center of shelf will continuously produce high flowers even when it will be displaced to the edge of the shelf.	True	Probably True	Insufficient Data	Probably False	False

Indicator 2: Recognition of Assumptions
Statement
The presence of soil bacteria and mycorrhizae can improve plant nutrition by making a certain amount of minerals which are available for plants. For example, many types of soil bacteria are involved in the nitrogen cycle, while mycorrhizal hyphae provides broader surface area for the absorption of nutrients, especially phosphate ions.

Assumption

The absence of mycorrhiza on the roots of plants effected the plants are not able to absorb nutrients. True False

Indicator 3: Deduction

Statement

Climate change can increase the growth of plant hoppers. The increase of plant hoppers can cause crop failure.

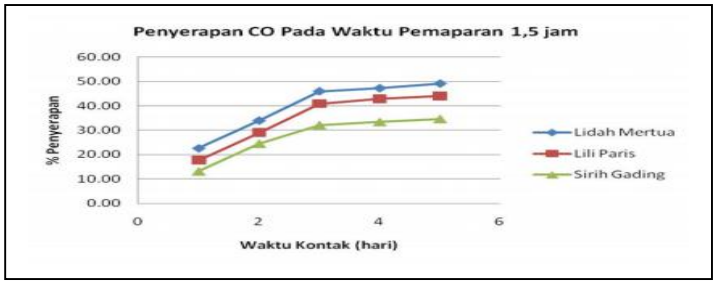
Conclusion

Planthopper population increase due to climate change True False

Indicator 4: Interpretation

Data

A study aims to determine the ability of decorative plants to absorb carbon monoxide, by using Sansevieria sp (tanaman lidah mertua), Spider plant (lili paris) and Scindapsus aureus (sirih gading) in an exposure time of 1.5 hours to produce a graph as follows.



Conclusion

The longer the contact time, the increase of the absorption is getting significantly. Yes No

Indicator 5: Evaluation of Arguments

Statement

A researcher examines the effect of fertilizer containing iron (Fe) on the plant growth. There are three groups and each group contained 10 plants. Group 1 uses fertilizer in sufficient quantities, group 2 uses less fertilizer, and group 3 has very much amount of fertilizer. The results showed that with plants little amount of fertilizer have the greatest growth. Thus, the researchers concluded that the fertilizer with iron content is better than fertilizer with nitrogen content. Strong Weak

The try out results show that small-scale test instrumentsthat have been developed is still not good. This is seen in the results of the test items validity analysis, which indicate that there are only 9 items classified as valid. The results of the analysis are then used as the basis of test items improvement in the revision stage. Furthermore, the revised test instruments areretried out on a large scale. The results of the analysis of large-scale trial are below, as follows.

Validity

The results of the validity analysis which used the Pearson correlation showed that there were 23 items classified as valid (sig <0.05). The 23 items consisted of 4 items as the indicator of the ability to make inferences, 6 items as the indicator of the ability to recognize assumptions, 6 items as the indicator of the ability of deductive reasoning, 4 items as the indicator of the

ability to interpret the arguments, 3 items as the indicator of the ability to evaluate arguments. The results of the 23 test items' validity analysis which were classified as valid can be seen in Table 1.2.

Table 1.2 Validity Analysis of Test items

Indicator	No. of Item test	Sig.
Making inference	1	0,015
	2	0,000
	3	0,035
	4	0,006
Identifying Assumption	5	0,000
	6	0,000
	7	0,000
	8	0,000
	9	0,009
	10	0,021
Deductive reasoning	11	0,018
	12	0,022
	13	0,005
	14	0,029
	15	0,004
	16	0,033
Interpreting argument	17	0,000
	18	0,000
	19	0,001
	20	0,000
Evaluating argument	21	0,000
	22	0,014
	23	0,001

The test items used in a multiple-choice test must be valid, it must be able to measure what to measure (Considine et al., 2005; Guidance, 2010). The results of the analysis showed that the validity of each indicatorof critical thinking skills had items classified as valid, so the 23 items were ready to be used for critical thinking skills test instruments for Biology students.

Reliability

The test items used in multiple-choice test must be reliable. By all means, it has to be consistent in measuring the same thing (Considine et al., 2005; Guidance, 2010). The results of the reliability analysis by using Cronbach's alpha indicated that the developed test instruments gotthe reliability value of 0.588. The reliability value indicated that the instruments classified in the category "quite reliable". Although the reliability of analysis results indicated that the developed test instruments got the high value of reliability (r> 0.7-1.0), but according to Watson and Glaser (2012), the test instruments with the reliability value less than 0.7 those instruments could already be implemented but on a limited scale. In addition, the instrument could also be used as an instrument for measuring the development of critical thinking skills. Based on the results of reliability analysis, the developed test instruments couldalready be used for critical thinking skills test instruments for Biology students.

Standard Deviation

The standard deviation analysis results of test items by using Pearson correlation with the reference category criteria Kolte (2015), showed that there were 10 items with different standard deviation in the excellent category (P> 0.35) and 13 items in good categories (0.20 <P> 0.35). The standard deviation analysis results of the test items can be seen in Table 1.3.

Table 1.3. The Standard Deviation Analysis of Test items

Indicator	No. Test Item	Standard Deviation (P)	Category
Making inference	1	0,256	Good
	2	0,393	Excellent
	3	0,223	Good
	4	0,286	Good
Identifying assumption	5	0,400	Excellent
	6	0,383	Excellent
	7	0,378	Excellent
	8	0,393	Excellent
	9	0,275	Good
	10	0,243	Good
Deductive reasoning	11	0,249	Good
	12	0,241	Good
	13	0,294	Good
	14	0,231	Good
	15	0,299	Good
	16	0,225	Good
Interpreting argument	17	0,464	Excellent
	18	0,565	Excellent
	19	0,357	Excellent
	20	0,412	Excellent
Evaluating argument	21	0,400	Excellent
	22	0,258	Good
	23	0,343	Good

The standard deviation analysis results of test items indicated that the test items which were valid had no poor quality in the standard deviation (P <0.2). Thus the test items with a value of P <0.2 were already acceptable and able to distinguish between students who have high ability and low ability (Mitra et al., 2009; Karelia dkk.2013). Based on the standard deviation analysis results of test items, it could be stated that the 23 items can already be used for critical thinking skills test instruments for Biology students.

Level of difficulty

The analysis results of the level of difficulty in the test items referring to the category criteria by Kolte (2015) indicated that there were 5 items classified in the category 'difficult' (p <0.3), 9 items were classified in the category of 'fair' (0.3 <p <0.7) , 9 items fall into 'easy' categories (p> 0.7). The analysis results of the level of difficulty in the test items can be seen in Table 1.4.

Table 1.4. The Analysis of Level of Difficulty in the Test Items

Indicator	No. test items	Difficulty index	Category
Making Inference	1	0,36	Fair
	2	0,46	Fair
	3	0,24	Difficult
	4	0,21	Difficult
Identifying assumption	5	0,46	Fair
	6	0,77	Easy
	7	0,94	Easy
	8	0,27	Difficult
	9	0,48	Fair
	10	0,40	Fair
Deductive Reasoning	11	0,97	Easy
	12	0,88	Easy
	13	0,96	Easy
	14	0,30	Fair
	15	0,11	Difficult
	16	0,71	Easy
Interpreting argument	17	0,37	Fair
	18	0,77	Easy
	19	0,50	Fair
	20	0,28	Difficult
Evaluating Argument	21	0,78	Easy
	22	0,32	Fair
	23	0,92	Easy

The analysis results of the level of difficulty in the test items showed that 23 items had a level of difficulty to the category of difficult, fair and easy. According to Boopathiraj and Chellamani (2013) items in a test should not be too difficult or too easy, so there must be a balance between those categories. Results of the analysis showed that the ratio of the level of difficulty in the test items was in the category of difficult, fairand easy is 5: 9: 9. Thus, it could be stated that the proportion is quite balanced, so it can be used for critical thinking skills test instruments for Biology students.

The results of the overall analysis in the test items can be concluded that there were 23 items that can be used for critical thinking skills test instruments for Biology students. However, this study had some limitations, so it needed to be re-examined at the next study. These limitations are;

1) The subjects of the try-out were limited only for the Biology students in State University of Malang. However, the accreditation of Biology program study in State University of Malang is A (very excellent), so that the results of the try-out are expected to be used to measure students' critical thinking skills in the majors biology / biology courses at other universities which have accreditation A and B. Furthermore, further research is expected to be tried out by involving Biology students in broader and wider universities.
2) The number of subjects in a large-scale try-out is limited only to 90 students. The next study is expected to increase the number of test samples. The more samples are used it will be a great opportunity to get a more accurate result (Sumanto, 2012).
3) The indicators of critical thinking skills that can be measured are limited to the indicators according to Watson and Glaser (2012). The instrument cannot measure indicators of critical thinking skills according to some other experts that have different indicators with Watson and Glaser (2012). For example the indicator to make a conclusion through induction by

Ennis (1985) and Facione (2000). However, Lai (2011) stated that in general, in terms of making a conclusion, it can be done inductively or deductively.

There were several limitations of the research development in critical thinking skills test instruments, but these studies had produced 23 items that were valid, reliable enough, the standard deviation that had been unacceptable, and the level of difficulty which was quite proportional. Test instruments referring to WGCTA patterns can be used as a test development of critical thinking skills, which is a test that can determine a student's strengths and weaknesses so that the results can be used as a basis for the development of critical thinking skills with appropriate learning activities (Watson and Glaser, 2012). In addition, the test instrument can also be used for research purposes (Wagner, 2002), particularly those which aimed at measuring the critical thinking skills of Biology students.

CONCLUSIONS

The development of critical thinking skills test instruments for Biology student which refer to the indicators and patterns WGCTA and the content in the form of general biological context, produced 23 valid items (Sig. <0.05) and quite reliable (0.588). The test instrument can be used as an instrument for the development of critical thinking skills in biology lectures, or for the benefit of research that aims to measure students' critical thinking skills in biology.

REFERENCES

Addy, T.M. & Stevenson, M.O. 2012. Implementation of Critical Thinking Exercises in Introductory Biology. *Proceedings of National Association of Biology Teachers 2012 Research Symposium*: NABT Four-Year College & University Section's.

American Association for the Advancement of Science. 2011. *Vision and Change in Undergraduate Biology Education a Call to Action*. Washington DC: Directorate for Education and Human Resources.

Ay, F.A., Karakaya, A. & Yimaz, K. 2015. Relations Between Self-Leadership And Critical Thinking Skills. *Procedia-Social and Behavioral Sciences.* 207: 29-41.

Boopathiraj & Chellamani. 2013. Analysis of Test Items on Difficulty Level and Discrimination Index in The Test For Research in Education. *International Journal of Social Science & Interdisciplinary Research*. 2(2): 189-193

Caroselli, M. 2009. *50 Activities for Developing Critical Thinking Skills*. Canada: Complimentary Resources from HRD Press.

Considine, J., Botti, M., Thomas, S. Design, Format, Validity, and Reliability of Multiple Choice Questions for Use in Nursing Research and Education. *Collegian*. 12 (1): 19-24

Dass, R. 2014. Literature and the 21st Century Learner. *Procedia-Social and Behavioral Sciences*. 123: 289-298.

Ejiogu, K.C., Yang, Z., Trent, J., & Rose, M. 2008. *Understanding the Relationship Between Critical Thinking and Job Performance Critical Thinking and Job Performance*. New Jearsey: Pearson.

Ennis, R.H. 1985. A Logical Basis for Measuring Critical Thinking Skills. *Educational Leadership.* 43(2): 44–48.

Facione, P.A. 2000. The Disposition Toward Critical Thinking: Its Character, Measurement, and Relation to Critical Thinking Skill. *Informal Logic*. 20 (1): 61–84.

Facione, P.A. 2011. *Critical Thinking : What It Is and Why It Counts*. Millbrae: Measured Reasons and The California Academic Press.

Fajrianthy, H.W. & Septarini, B.G. 2016. Developing Critical Thinking Test Utilising Item Response Theory. *Jurnal Penelitian dan Evaluasi Pendidikan*. 20 (1): 45-55

Fong, L.L., Sidhu, G.K., & Fook, C.Y. 2014. Exploring 21st Century Skills among Postgraduates in Malaysia. *Procedia-Social and Behavioral Sciences*. 123:130-138.

Friedel, C., Tech, V. & Eckhardt, E. 2008. Overtly Teaching Critical Thinking and Inquiry-Based Learning: A Comparison of Two Undergraduate Biotechnology Classes. *JournaL of Agricultural Education.* 49(1): 72-84.

Guidance, 2010. *Multiple-Choice Testing as an Assessment Tool*. (Online) (www.wales.gov.uk/publications), diakses tanggal 28 Agustus 2016.

Halpern, D.F. 1999. Teaching for Critical Thinking: Helping College Students Develop the Skills and Dispositions of a Critical Thinker. *New Directions For Teaching And Learning*. 80: 69-74

Hambleton, R.K., & Jones, R.W. 1993. An NCME Instructional Module on: Comparison of Classical Test Theory and Item Response Theory and Their Applications to Test Development. *Educational Measurement: Issues and Practice.* 12(3). 38-47.

Husband, G. 2006. *an Analysis of Critical Thinking Skills in Computer Information Technology Using California Critical Thingking Skills Test*. Menomonie: University of Wisconsin-Stout.

Jerald, C.D. 2009. *Defining a 21st Century Education*. USA: For the Center for Public Education.

Karelia, B.N., Pillai, A., & Vegada, B.N. 2013. The Levels of Difficulty And Discrimination Indices and Relationship between Them in Four-Response Type Multiple Choice Questions of Pharmacology Summative Tests of Year II M.B.B.S Students. *IeJSME*.7(2): 41-46

Kolte, V. 2015. Item Analysis of Multiple Choice Questions in Physiology Examination. *Indian Journal of Basic and Applied Medical Research*. 4(4): 320-326

Lai, E.R. 2011. *Critical Thinking: A Literatur Review*. New Jearsey: Pearson.

Mitra, N.K., Nagaraja, H.S., Ponnudurai, G., & Judson, J.P. 2009. The Levels of Difficulty and Discrimination Indices in Type A Multiple Choice Questions Of Pre-clinical Semester 1 Multidisciplinary Summative Tests. *IeJSME*. 3(1): 2-7

Musa, F., Mufti, N., Latiff, R.A., & Amin, M.M. 2012. Project-based Learning (PjBL): Inculcating Soft Skills in 21st Century Workplace. *Procedia-Social and Behavioral Sciences*. 59: 565-573.

NCREL & Meitri Group. 2003. *enGauge 21st Century Skills: Literacy in the Digital Age*, (Online), (http://www.ncrel.org/engauge.org/engauge), diakses pada tanggal 28 Agustus 2016.

Osman, K., & Marimuthu, N. 2010. Setting new learning targets for the 21st century science education in Malaysia. *Procedia-Social and Behavioral Sciences.* 2(2): 3737-3741.

Osman, K., Hiong L.C., & Vebrianto, R. 2012. 21st Century Biology: An Interdisciplinary Approach of Biology, Technology, Engineering and Mathematics Education. *Procedia-Social and Behavioral Sciences*. 102:188-194.

Pheeraphan, N. 2015. Enhancement of the 21st Century Skills for Thai Higher Education by Integration of ICT in Classroom. *Procedia-Social and Behavioral* Sciences.103: 365-373.

Rodzalan, S.A., & Saat, M.M. 2015. The Perception of Critical Thinking and Problem Solving Skill among Malaysian Undergraduate Students. *Procedia-Social and Behavioral Sciences*: 172: 725-732.

Sahin, M.C. 2009. Instructional Design Principles for 21st Century Learning Skills. *Procedia-Social and Behavioral* Sciences. 1(1): 1464-1468.

Slameto. 2014. Developing Critical Thinking Skills through School Teacher Training "Training and Development Personnel" Model and Their Determinants of Success. *International Journal of Information and Education Technology*. 4(2): 161-166.

Stassen, M.L.A., Herrington, A., & Henderson, L. 2011. *To Improve The Academy*. San Francisco, CA: Jossey-Bass Publishers.

Sumanto, D. 2012. Precision and Accuracy of Results of Quantitative Research Based on The Random Sampling Determination. *Jurnal Litbang Universitas Muhammadiyah Semarang*. Page: 45-53

The Partnership for the 21st Century Skills, 2007. *Framework for 21st Century Learning*. (Online) (http://www.p21.org/our-work/p21-framework), diakses tanggal 28 Agustus 2016.

The Partnership for the 21st Century Skills. 2009. *Framework for 21st Century Learning* . (Online) (http://www.p21.org/our-work/p21-framework), diakses tanggal 28 Agustus 2016.

Thompson, C. 2011. Critical Thinking across the Curriculum : Process over Output. *International Journal of Humanities and Social Science*.1(9): 1-7.

Turiman, P., Omar J., Daud, A.M., & Osma, K. 2012. Fostering the 21st Century Skills through Scientific Literacy and Science Process Skills. *Procedia-Social and Behavioral Sciences*. 59: 110-116.

Wagner, T.A. 2002. Critical thinking: The development of A New Measure. *Thesis*. Blacksburg, Virginia: Virginia Polytechnic Institute and State University.

Watson, G. & Glaser, E.M. 2012. *Watson-Glaser™ Critical Thinking Appraisal User-Guide and Technical Manual*. UK: Pearson Education Ltd.