Research and Evaluation in Education Journal
Volume 1, Number 2, December 2015 (pages 186-198)
Available online at: http://journal.uny.ac.id/index.php/reid

A FORMATIVE ASSESSMENT MODEL OF CRITICAL THINKING IN MATHEMATICS LEARNING IN JUNIOR HIGH SCHOOL

[1)]R. Rosnawati; [2)]Badrun Kartowagiran; [3)]Jailani
[1)2)3)]Yogyakarta State University, Indonesia
[1)]rosnawati@uny.ac.id; [2)]badrun_kartowagiran@uny.ac.id; [3)]jailani@uny.ac.id

Abstract

This study aims to obtain a valid and reliable formative evaluation model of critical thinking. The method used in this research was research and development by integrating Borg & Gall's model and Plomp's development model. The ten steps of Borg & Gall's model were modified into five stages as the stages in the Plomp's model. The subjects in this study were 1,446 students of junior high schools in Yogyakarta Special Region or *Daerah Istimewa Yogyakarta* (DIY), 14 mathematics teachers, and six experts. The content validity employed was expert judgment, the empirical validity and reliability used were loading factor, item analysis using partial credit model (PCM) 1PL, and the relationship between disposition and critical thinking skill used was structural equation modeling (SEM). There are five aspects of critical thinking skill: mathematic reasoning, interpretation, analysis, evaluation, and inference, which entirely composed of 42 items. The validity of the critical thinking skill instruments achieves a significant degree as indicated by the lowest and highest loading factors of 0.38 and 0.74 subsequently, the reliability of every aspect in a good category. The average level of difficulty is 0.00 with the standard deviation of 0.45 which is in a good category. The peer assessment questionnaire of critical thinking disposition consists of seven aspects: truth-seeking, open-minded, analysis, systematic, self-confidence, inquisitiveness, and maturity with 23 items. The critical thinking disposition validity achieves the significant degree as indicated by the lowest and the highest factor loading of 0.66 and 0.76 subsequently, and the reliability of every aspect in a good category. Based on the analysis of the structural equation model, the model fits the data.

Keywords: *formative assessment, critical thinking, mathematics learning*

Introduction

Critical thinking is one of the most important skills in thinking which must be owned by students because critical thinking will make someone easier to process and use information found to solve any problem. In general, critical thinking is a mental activity to evaluate certain things by using rational, systematic, and reflective reasons with an emphasis on making decisions about what to believe and do. Ennis (1985, p.44; 1993, p.179; 2011, p.10) states that 'critical thinking is a reasonable reflective thinking focused on deciding what to believe or do.' According to Ennis, skills associated with critical thinking can be learned and transferred from one to other disciplines. This definition of critical thinking incorporates critical thinking skills and critical thinking disposition.

On the contrary, McPeck (Mason, 2007, p.340) states that critical thinking is specific, cannot be taught freely in a particular field. To become a critical thinker in mathematics, it would be very difficult if a person only has a little knowledge about the field. McPeck's definition does not distinguish critical thinking skills and critical thinking disposition.

Siegel (Mason, 2007, p.342) emphasizes the concept of a strong relationship between critical thinking with rationality and defines critical thinking as follows: critical thinking means to be 'appropriately moved by reasons', and to be rational is to 'believe and act on the basis of reasons'. Siegel's concept (1985, p.78) about critical thinking keeps the assessment component of reasoning (reason assessment component) on the domain of skills and critical attitude component on the disposition domain. Siegel (2010, p.140) also emphasizes the importance of disposition: 'the disposition to engage in and be guided by assessment'.

Thinking disposition is a tendency towards certain intellectual behavior patterns. Nieto and Saiz (2011, p.206) state that critical thinking disposition becomes a consistent internal motivation to be involved in issues and decisions using critical thinking. Critical thinking disposition can also be defined as an individual gate to reach critical thinking activities (Irani, et al., 2007, p.3). In the same spirit, Facione, et al. (2000, p.4) claim that thinking disposition is a constellation of attitudes, intellectual virtues, and habits of critical thinking. In other words, critical thinking disposition is a constellation of attitudes and intellectual virtues which becomes a consistent internal motivation used to achieve critical thinking activities.

Critical thinking disposition affects the success not only in teaching-learning and other things, but also in controlling ourselves and approaching other people (Kawashima, et al, 2007, p.188). Someone who has a critical component of disposition will have certain characteristics and certain skills. Disposition is a character that underlies decisions and actions, a character that is not arbitrary and a character of a person to objectively assess the relevant facts. Critical thinking disposition serves as intellectual honesty, fairness, sympathy, and objectivity. Therefore, the skills and principles of critical thinking disposition are essential for critical thinkers.

Critical thinking in mathematics actually has been a general requirement in the curriculum of junior high school mathematics. To obtain accurate information about the strengths and weaknesses of the students, critical thinking is conducted systematically by collecting, formulating, and processing data. Formative evaluation is needed for the facts used to identify the needs of students with the aim of achieving the critical thinking skills of a class or group. Mardapi (2008, p.9) says that brief evaluation can be defined as the process of gathering information for the achievement in the study of a class or group.

There is no consensus about how critical thinking should be measured. Results of previous studies show there are a number of instruments of critical thinking, for instance: Watson-Glaser Critical Thinking Appraisal Skills (WGCTA, Watson & Glaser, 1980, pp.1-13), Ennis-Weir Critical Thinking Essay Test (EWCTET, Ennis & Weir, 1985, pp.1-14), Cornell Critical Thinking Test (CCTT, Ennis, Millman, & Tomko, 1985, pp.1-18), the California Critical Thinking Skills Test (CCTST, Facione, 1990, pp.1-84), California Critical Thinking Disposition Inventory (CTDI, Facione & Facione, 1995,

pp.1-5), and Halpen Critical Thinking Using Everyday Situation Assessment (HCTAES, Halpern, 2010, pp.1-34). The instrument was developed based on a common problem. The type of instrument developed varies and is not specific to mathematics in junior high school.

The afore-mentioned description illustrates that the formative evaluation of students' critical thinking has an important role for the development of education quality. Considering the importance of formative assessment of critical thinking, its development becomes an urgent requirement. The purpose of this study is to design a formative evaluation model of critical thinking. The main problem of this research is how to develop a valid and reliable formative evaluation model of critical thinking in teaching-learning mathematics.

Research Method

In accordance with the objective to be achieved in this study, research and development was employed. The product of this research is in the form of a formative evaluation model of critical thinking in teaching-learning process in junior high school's mathematics that contains critical thinking construct, critical thinking skills instruments, scoring rubric, critical thinking disposition questionnaire, and guidelines for the use of formative evaluation of critical thinking.

This study employed a modified research and development model of Borg & Gall combined with Plomp's development model. The modification mentioned in this research is a combination between Borg & Gall development model and Plomp's. Of the ten steps of Borg & Gall's model, they are summarized the model into five stages as in Plomp's (2007). The steps which were taken were (1) conducting a preliminary study to analyze the product which will be developed, (2) developing a design (3) developing the product and validation of the product, (4) a field trial, (5) dissemination/ implementation.

The subjects of this research were grouped into the subjects of development stage and the subjects of main field trial. The subjects of development stage were engaged in the development instrument model of critical thinking, while the subjects of trial stage were those engaged in the field.

The experts of the validation were experts of mathematics education, judgment experts in education, and psychologists. The subjects of the pilot trial were 656 eighth graders of junior high schools in Special Region of Yogyakarta -- or *Daerah Istimewa Yogyakarta* (DIY) and four mathematics teachers, while the subjects of the main field trial were 810 eighth graders of junior high schools in DIY and eight mathematics teachers.

Data analysis techniques employed in this study were qualitative and quantitative data analysis techniques. Qualitative data analysis was the analysis of the data obtained from experts' judgments and practitioners as well as other research subjects that provided input. On the other hand, quantitative data analysis techniques used in this study included validity and reliability analysis techniques and the analysis of test item by using PCM 1PL approach and goodness of fit model.

To prove the content validity, the research was conducted based on the statement from an expert (Mardapi, 2008, p.18) in a research using Delphi. Construct Validity is proven by a loading factor which was derived from factor analysis technique. An observable variable is declared valid to measure latent variables when the value of loading factor is greater than 0.3 (Fernandes, 1984, p.28; Schumacher & Lomax, 2004, p.176).

In order to test the hypotheses about the existence of the constructs or to seek the constructs in the variables, factor analysis was used. Factor analysis is the study of the interdependence between variables to find a new set of variables which is considerably less than the original variables and to indicate which original variables are common divisors (Basilevsky, 1983, p.98). In addition, factor analysis can be used to inform validity (Thompson, 2004, p.4).

Besides, reliability analysis technique tested the consistency of the data measured which was generated by the instruments; the value of reliability coefficient indicated the level of instruments consistency, higher level of reliability, and smaller possibility of measurement error. The reliability analysis

technique used was multidimensional reliability by using construct reliability (CR). This technique was chosen because the instruments were developed containing several aspects. Kartowagiran (2008, p.181) mentions that multidimensional factor on the test analyzed by unidimensional models causes improper estimation of capabilities and provides misleading information.

Rasch's model for polytomous items was used to analyze the test item. According to Wright & Masters, partial credit model (PCM) is also appropriate to analyze the response to the judgment of critical thinking and conceptual understanding in science (Van der Linden & Hambleton, 1997, pp.101-102). PCM 1PL model analyzes the data focusing on the difficulty level parameters. The item is said to be good if the difficulty index is more than -2.0 or less than 2.0 (Hambleton & Swaminathan, 1985, p.36). The test item analysis was done using QUEST software.

The test on the model was done by using Structural Equation Model analysis technique or SEM. SEM is the analysis technique of latent variables, indicator variables, and measurement errors directly to determine the relationship between the latent variables with other variables, as well as the measurement error. SEM is a confirmatory analysis of the variables. The focus of the analysis in this section is to examine the relationship between the model of dependent and independent constructs (Ghozali & Fuad, 2008, p.3; Kline & Klammer, 2001, p.213).

The test criteria used some of this model's indicators, namely (1) Chi-square; (2) a comparison between Chi-square with degree of freedom; (3) NCP or non-centrality parameter; (4) Root Mean Square Error of Approximation (RMSEA); (5) Expected Cross Validation Index (ECVI); (6) AIC and CAIC; (7) NFI or Normed Fit Index; (H) NNFI or Non-normed Fit Index; (8) Comparative Fit Index or CFI; (9) IFI or Incremental Fit Index, and (10) Relative Fit Index or RFI (Joreskog & Sorbom, 1996, p. 124; Hair, et al., 2010, p.22). Thus, the criteria employed in this study were Chi-square, a comparison between Chi-square with degree of freedom, RMSEA, NFI, NNFI, CFI, IFI, and RFI.

Findings and Discussion

Findings

The developed formative evaluation model is procedural model. Development is done by reviewing the theory of critical thinking and the model's development, and testing a construct model of critical thinking.

Formative Evaluation Development of Critical Thinking in Mathematics Teaching-Learning Process in Junior High School

To conduct formative evaluation of critical thinking in mathematics teaching-learning process for junior high school students, it is necessary to gather information related to the cognitive development of junior high school students in order to obtain real expectation related to their ability to think.

Classic Piaget's theory is focused on the intellectual development of children. According to Piaget, junior high school students (ages 11-16) are at the formal operational stage (Hamilton & Ghatala, 1994, p.218; Bjorklund, 2005, p.85; Siegler & Alibali, 2005, p.53). The ability measurement index at the formal operational stage is hypothetico-deductive reasoning: the ability to think scientifically and to apply the scientific method to cognitive tasks. At this stage, they have the ability to draw conclusions based on not only physical facts but also hypotheses, to hold their opinions and explain the opinions, and to have the ability to make inferences from general to specific situations (deductive reasoning). In addition, to develop deductive reasoning skills, Piaget theorizes that children at the formal operational stage are able to think inductively to conduct special observations and makes generalizations (Bjorklund, 2005, p.100). Neo-Piaget's theory of Fischer states that junior high school students have the ability to think abstractly (Fischer, 1980, pp. 491-493; Bjorklund, 2005, p.107). In other words, junior high school students are able to develop their critical thinking skills.

To obtain accurate information about students' strengths and weaknesses related to critical thinking skills, formative evaluation of critical thinking, especially in mathematics, is needed. This information can be used to

identify students' needs with the aim of achieving critical thinking skills to think or group classes. Junior high school mathematics contains some materials, such as numbers, algebra, geometry and measurement, as well as statistics and probability. Based on the previous theory, critical thinking is subject-specific in which the evaluation conducted is very specific to the particular content.

In order to become a critical thinker in algebra content, it will be very difficult if a person only has a little knowledge about algebra. It is similar for numbers, geometry and measurement, as well as statistics and probability. Thinking in each sphere in mathematics is very specific. Therefore, the collection of information related to critical thinking skills can be carried out based on specific content by considering the thinking on any content.

Thinking on numbers is thinking related to numbers (number sense) and counting the numbers (quantitative) involving operations on numbers. The operations performed by students at the level of numbers tend to not involve the relational aspects of the operations, but only focus on the problem of counting (Kieran, 2004, p.140). Algebraic thinking is associated with the generalization of arithmetic operations and operation at the variables. Van de Walle, et al. (2011, p.262) state that algebraic thinking or reasoning involves the generalization of experience with numbers and computing, idea formulation of using symbols, and exploration of the concept of patterns and functions. Geometric thinking requires the ability to make a proper connection between the definition of the concept of geometry and the concept of images/shapes (Crowley, 1987, p.1). Statistical thinking involves how someone understands the nature of sampling, how to make inferences from the sample of population, how someone understands random phenomena, how data are generated to estimate probabilities, and how, when, and why the existing inference tool can be used to assist investigation process (Ben-Zvi & Garfield, 2005, p.7).

Critical thinking disposition is a component of critical thinking that becomes a consistent internal motivation involved in issues and decisions using critical thinking. The concept of critical thinking is as follows.

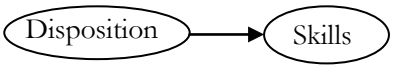

Figure 1. The Relationship between the Dimension of Disposition and Skills

The Construct Development of Critical Thinking Skills in the Contents of Algebra

This research develops critical thinking instruments in algebra contents, while the topics in junior high school mathematics are pattern, algebra operations, sets, relations and functions, and the system of linear equations. Based on the critical thinking study and contents of Algebra mentioned before, this study identifies the conceptualization or the type of critical thinking in algebra topics, namely: 1) using reasoning in algebra problems, 2) providing examples related to the concepts in algebra (interpretation), 3) checking arguments or statements related to algebra (analysis), 4) providing an assessment of two statements or concepts in algebra (evaluation), 5) using the concepts to draw conclusions and make generalizations (inference), and 6) declaring the results of reasoning to justify the reasoning (explanation). The study is used as the basis for formulating indicators of critical thinking skills.

The test result of the CFA first-order measurement model with 42 items showed p-value = 0.091 ($p > 0.05$) and RMSEA = 0.017 (RMSEA <0.08). In other words, the 42 items were valid indicators for the measurement of critical thinking skills in algebra contents. Thus, the measuring instruments for critical thinking skills met the assumption of unidimensionality so that the model application with PCM 1PL approach could be conducted.

The test result indicated that there was a change in the measurement construct of critical thinking skills. This change occurred because there was no explanation indicator so that the six aspects of critical thinking skills turned into five aspects. Based on the output of rotated factor matrix, there are five factors of critical thinking. They are mathematical reasoning as the 1st factor, interpretation as the 2nd factor, analysis as the 3rd factors,

evaluation as the 4th factor, and inference as the 5th factor. Furthermore, the distribution of the items of critical thinking skills instruments in the contents of algebra was served. The forty-two items were included in the 1st factor to 5th factor with the total variant explained as much as 51.39%. It showed that the five factors developed could measure the variables of critical thinking skills as much as 51.39%.

Second Order CFA model testing on the 42 items resulted in p-value = 0.0627 and RMSEA = 0.019. Therefore, it can be concluded that the model was in 'fit' category.

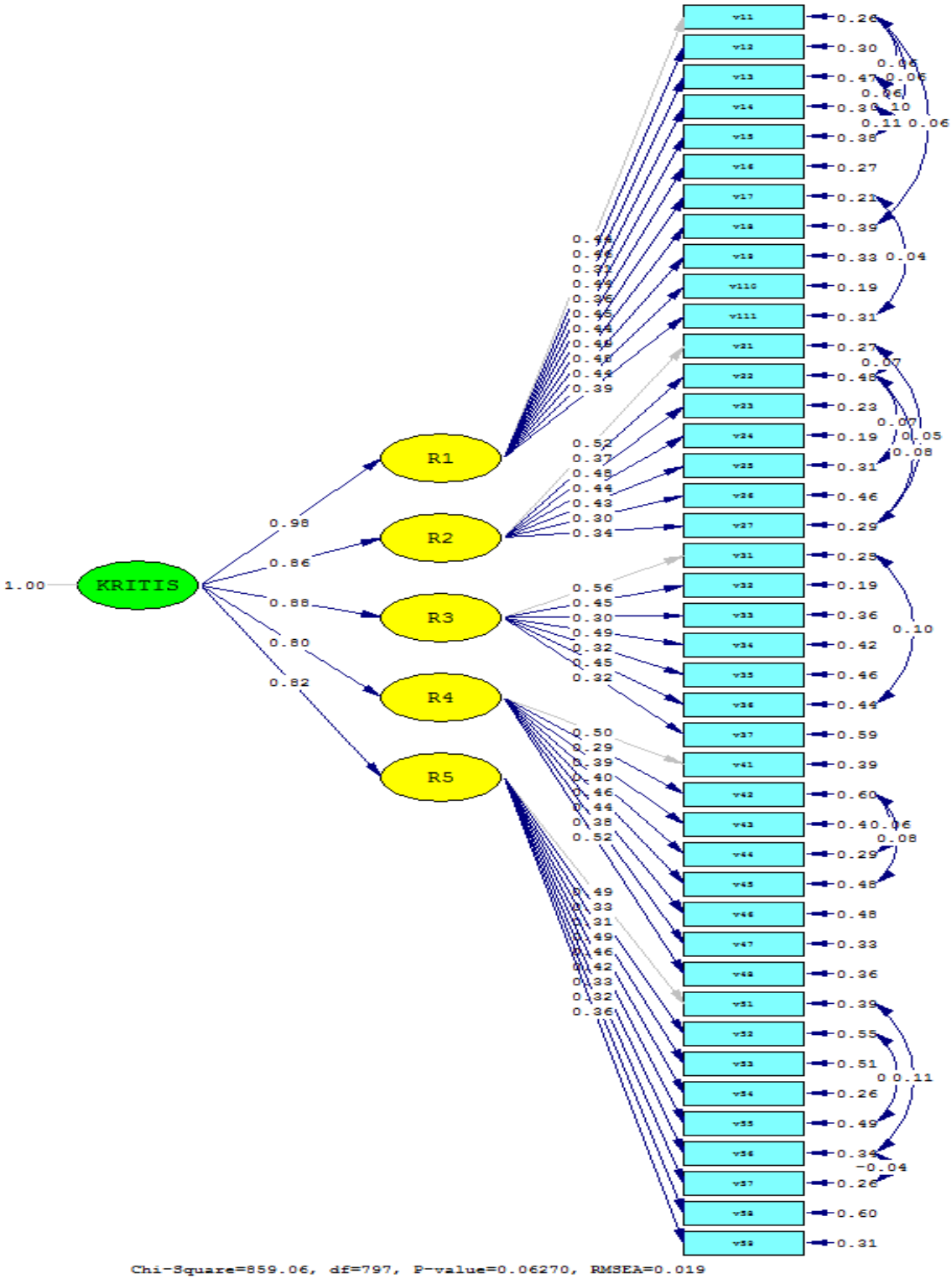

Figure 2. Second Order CFA of Critical Thinking Skills

Item analysis was conducted by the goodness of fit testing which was done to test critical thinking skills or per item. The overall testing of fit tests used the rules which were developed by Adams and Khoo (1996, p.30), based on INFIT Mean of Square mean value (Mean INFITMNSQ) and standard deviation or observing the mean value of INFIT t (Mean INFIT t) and its standard deviation. If the mean of INFIT MNSQ is about 1.0 and 0.0 for standard deviation or mean INFIT t almost 0.0 and 1.0 for the standard deviation, the entire test fits with PCM 1 PL model. Based on analysis, the mean value of INFIT MNSQ 1.06 (about 1) and standard deviation of 0.40 (about 0.0), then the entire test is fit with PCM 1PL model.

Testing on the determination of fit in every item on the model followed the rules of Adams and Khoo (1996, p.30), which state that an item is fit to the model if the value of INFIT MNSQ is between 0.77 and 1.33. The MNSQ INFIT value is between 0.79 to 1.24. With the acceptance limit of the item using INFIT MNSQ or fit according to the model (between 0.77 to 1.30), all items, 42 items, were fit with PCM 1PL. Based on the analysis of the difficulty index, the items lied between -1.08 to 0.69 with an average of 0.0 and standard deviation of 0.45. Hence, based on difficulty, all of the 42 items were fit. The difficulty index for the aspects of mathematical reasoning, interpretation, analysis, evaluation, and inference was 0.082, -0.262, -0.191, 0.073, and 0.186. The inference aspect had the highest difficulty index which was compared to other aspects.

Feedback

To facilitate the provision of information about students' abilities and weaknesses from the result of the test on critical thinking skills, a description of the band-scale characteristics were developed as listed in the following table. Band-scale was developed by identifying items that represented a certain score range, then a description was made on the characteristics which were represented by the set of items in each range of scores.

Table 1. The Band-scale Characteristics of Information about Students' Abilities and Weaknesses from the Result of the Test on Critical Thinking Skills

Low (0-25)	High (50-75)	Medium (25-50)	Very High (75-100)
Students have knowledge to recognize the problem and provide examples to explain ideas.	Students can provide reasons with information, give assessments, examine ideas and detect arguments.	Students can provide reasons with information, use inductive and deductive proof, and provide an assessment of the two statements.	Students can provide reasons with information, give assessments, examine ideas and detect arguments, draw conclusions, and make generalizations.
Students can understand the basic arithmetic operations, patterns, and simple algebraic relations.	Students can understand the basic arithmetic operations, patterns, and simple algebraic relations, read and interpret tables and graphs, represent algebraic models, provide an assessment of the two statements, and provide an analysis of the relationship between two or more concepts.	Students can understand the basic arithmetic operations, binary operations, patterns, and simple algebraic relations, read and interpret tables and graphs, and represent algebraic models.	Students can understand the basic arithmetic operations, patterns, and simple algebraic relations, read and interpret tables and graphs, represent algebraic models, provide an assessment of the two statements, provide an analysis of the relationship between two or more concepts, draw both inductive and deductive conclusions, and make generalization.

By using the design of the norm measuring instruments of critical thinking skills based on the reference domain, the users of the instruments would be facilitated to describe the critical thinking skills of the test participants in learning mathematics. The description of critical thinking skill of the test's participants could ease the users to overcome the weaknesses of the participant in critical thinking skills, especially in the topic of algebra.

The Development of Critical Thinking Disposition Constructs

Disposition is a marker for the significance of critical thinking and critical thinking disposition that can be obtained by a person slowly over time. A person with low levels of a disposition will have no worries while facing a complex question, searching for different solutions to a problem, and questioning different solution of a problem. It will be different from a person who has a high disposition.

Facione (2000, p.4) states there are seven aspects of critical thinking disposition: truth-seeking, open-mindedness, analytic aspect, systematic aspect, confidence, inquisitiveness, and maturity. *Truth-seeking* is a character that always someone wants to find the best knowledge in a particular context, is courageous to ask questions, and is honest and objective in pursuing the investigation even if the findings do not support an individual's personal interests or preconceived opinions. Students who have understood certain knowledge will recognize the existence of a set of facts or relevant reasons. Hence, in facing a problem, people who have a disposition will tend to adjust their belief in accordance with the facts and reasons. The search for truth is one way to give consideration of additional facts, reasons, or perspectives to one's mind when dealing with a problem.

Open-mindedness is a character that someone is tolerant towards different views and is sensitive to the possibility of its own mistakes. This character is very important for students who currently face a pluralistic - multi-cultural society. This skill is important for students to understand what other people think and the possibility of mistakes they can make.

Analytic is a character that someone upholds the application of reasoning and the use of evidence to resolve a problem, anticipates potential conceptual or practical difficulties, and is consistent. Being analytic is the core of disposition for an inquiring mind. People with these characteristics tend to anticipate the consequences of an event or an idea, and use the reasons of several strategies to address an issue.

Systematical is a character of being organized, orderly, focused, and diligent in an investigation. Organized approach to problem solving and decision making is a special quality of wise people regardless of the domain of the problem addressed.

Self-Confidence can be in the process of self-reasoning. Confidence in critical think-ing allows a person to trust the power of reason for self-assessment and lead others in looking at the issue rationally. The appropriate level of confidence in critical thinking is an increase in relation to the maturity of a person and in relation to one's critical thinking skill mastery. It is certainly going to be a path of desirable development for all students.

Inquisitiveness in critical thinking is a character of one's intellectual curiosity and desire to learn knowledge, even when the application of the knowledge is not visible. Inquisitiveness can be in the form of intellectual inquisitiveness and inquisitive-ness in the limitation of the characteristics of various educational backgrounds.

Maturity targets the disposition to be wise in one's decision-making. One's matur-ity can be characterized as a structured approach of a problem, investigation, and in decision–making. Some situations recognize more than one reasonable choice. Other conditions demand that the assessment must be made many times based on the standard, context, and evidence that block the certainty. This maturity can be seen when a person responds to a problem wisely and with structured attitude and when he/she makes complex decisions that involve a variety of interests. Those seven studies were used for

the preparation of indicators disposition of critical thinking.

Someone who has a component of critical attitude will have a certain character that is as good as certain skills: a character where someone is likely to find out something that underlies a decision and an action; a character that is not impartial and not arbitrary; a character that assesses relevant facts objectively; a character that has the values of critical thinking aspect, such as intellectual honesty, fairness, sympathy, and objectivity. Therefore, the skills and principles of critical thinking are essential for critical thinkers.

Critical thinking disposition instruments used in the study were in the form of peer assessment questionnaire. The peer assessment format was chosen because peer assessment in junior high schools indicates an assessment that has a positive effect on students' satisfaction and learning effectiveness in different disciplines (Cheng, et al., 2008, p.38). It is also stated by Hammond (2013, p.48) that peer assessment gives a positive effect on increasing motivation and confidence. A similar opinion is expressed by Falchikov (2005, p.57) that peer assess-ment deepens students' understanding of their own learning and empowers students to become more actively engaged and self-directed in their learning process.

The test result of the CFA first-order measurement model with 23 items showed p-value = 0.089 (p> 0.05) and RMSEA = 0.025 (RMSEA <0.08). In other words, the 23 items were valid indicators for the measurement of critical thinking disposition. The twenty-three items were included in the 1st factor to 7th factor with the total variant explained as much as 70.98%. It showed that the seven factors developed could measure the variables of critical thinking skills as much as 70.98%

CFA Second Order CFA model testing on the 23 items resulted in p-value = 0.0727 and RMSEA = 0.023. Therefore, it could be concluded that the model could be included in fit category. It showed that this model fitted to the data. In other words, the 23 items were valid indicators for measuring the critical thinking disposition construct.

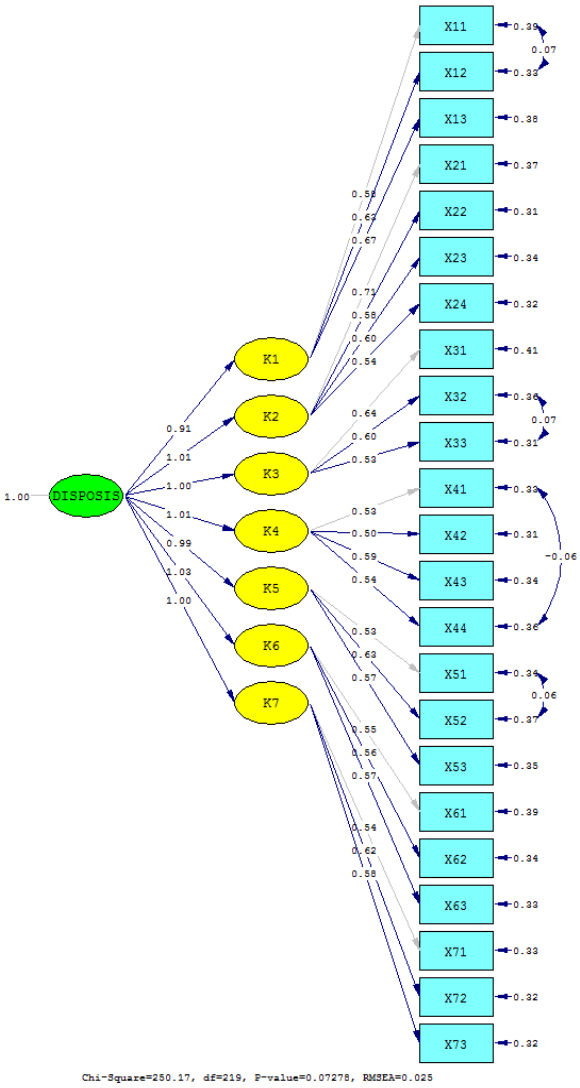

Figure 3. Second Order CFA of Critical Thinking Disposition

Discussion

The result of qualitative analysis of the formative evaluation model has a good qualification. Thus, validation through the examination process by the experts has led to the formulation of instruments in accord-ance with the indicators. Mardapi (2008, p.15) says that validation, both theoretical and empirical, is an important step.

The result of testing instruments for critical thinking skills showed that based on the main field trial, of forty-two items of instruments, they were all valid because they had a loading factor above the value of 0.3 and the value of t which was greater than 1.96. The five aspects of the critical skill

instruments had construct reliability above 0.7 so that the instruments in all five aspects were reliable. p-value = 0.0627 (> 0.05), the value of RMSEA = 0.019 (<0.80) and the value of GFI = 0.97 (> 0.90) fulfilled the requirement as fit model. The test result provided an indication that all requirements were met as a fit model. Thus, the measurement instruments can already be used as a measurement model of critical thinking skills.

The result of testing instruments for critical thinking disposition showed that the 23 items of critical thinking disposition instruments, all of them were entirely valid because they had a loading factor above the value of 0.3 and the value of t which was greater than 1.96. The criteria of hypothetic-al measurement for the goodness of fit model with empirical data showed p-value=0.073 (> 0.05), the value of RMSEA=0.025 (<0.80) that qualified as fit model, and the value of GFI = 0.91 (> 0.90) also qualified as fit model. Hence, the measurement instruments of critical thinking disposition can already be used as a measurement model that is suited to collect data about students' critical thinking disposition.

Quantitative test result, for both the instruments of critical thinking skills and disposition analyzed by using EFA and CFA, shows the items of the instruments developed are valid and reliable items. For the instruments of critical thinking skills, based on the analysis of difficulty index, the items lie between -1.08 to 0.69 with the average of 0.0 and the standard deviation of 0.42. Thus, based on difficulty, all of the 42 items are fit. The difficulty index for each aspect is in the range of -0.262 to 0.186. The inference aspect has the highest difficulty index compared to other aspects.

The Result of the Goodness of Fit Model

The report on the result of the goodness of fit model is the result of analysis of 395 fit data. The result of the goodness of fit model with empirical data can be seen in figure 4.

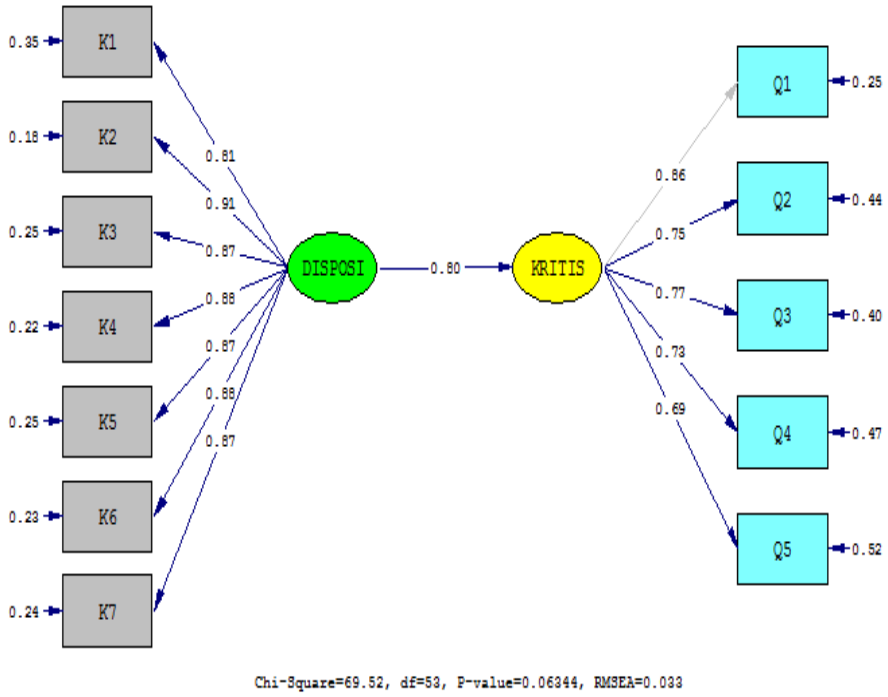

Figure 4. SEM on the Critical Thinking Construct

Based on the analysis of structural equation model as shown in Figure 4, it is shown that the Chi-square = 69.52, df = 53, p-value = 0.063 (>0.05), and RMSE = 0.033 (<0.08), NFI=0.99, NNFI=1.00, CFI=1.00, IFI=1.00, and RFI=1.00 The result shows that the model is fit with the data. After obtaining the fit model, the coefficient of

determination value (R^2) of reduced form equation from the variables observed with their latent variables can be seen. The coefficient of determination value of each variable is explained in Table 2.

Table 2. The Result of SEM on the Critical Thinking Construct

No	Variable	Loading Factor	*t*-value	R^2
Critical Thinking Skills				
1	Mathematical Reasoning	0.86	-	0.75
2	Interpretation	0.75	14.67	0.56
3	Analysis	0.77	15.63	0.60
4	Evaluation	0.73	14.66	0.53
5	Inference	0.69	13.22	0.48
Critical Thinking Disposition				
1	Truth-seeking	0.81	16.60	0.65
2	Open-mindedness	0.91	19.92	0.82
3	Analytic	0.87	18.59	0.75
4	Systematic	0.88	19.16	0.78
5	Self-Confident	0.87	18.54	0.75
6	Inquisitiveness	0.88	18.90	0.77
7	Maturity	0.87	18.63	0.76

The aspect that affects critical thinking skills the most is mathematical reasoning, with the value of $R^2 = 0.75$ which means that the aspect of mathematical reasoning contributes 75% towards critical thinking. The aspect that affects critical disposition the most is open-mindedness with the value of $R^2 = 0.82$ which means the aspect of open-mindedness contributes 82% to critical thinking disposition. γ parameter is the estimation of the direct influence of latent exogenous variable towards endogenous variable of critical thinking disposition. The estimated value of γ parameter is 0.81 as the direct effect of critical thinking disposition skills towards critical thinking skills. The significance of γ parameter is at 13.77 ($t > 1.96$) with the value of $R^2 = 0.64$. It means that critical thinking disposition skills contribute as much as 65% towards critical thinking skills, and the rest (35%) is influenced by other variables.

Conclusion and Suggestions

Conclusion

Formative assessment model development of critical thinking begins with the development of instruments of critical thinking skills, and critical thinking disposition instruments.

The instruments of critical thinking skills in the content of algebra are in the form of multiple choice with four options and with open reason developed by including five aspects, namely: mathematical reasoning, interpretation, analysis, evaluation, and inference, and are proven to fit with the data. When the forty-two items based on the analysis of PCM 1PL are viewed from the average index of difficulty, they are qualified as the instruments of critical thinking skills.

The instruments of critical thinking disposition in the form of Likert scale questionnaire that contain the purposes of the aspects, truth-seeking, open-mindedness, analytic aspect, systematic aspect, self-confidence, inquisitiveness, and maturity have good qualifications.

Based on the analysis of structural equation model, critical thinking disposition affects critical thinking skills. The significance test of critical thinking skills and critical thinking disposition obtained from the data shows that critical thinking disposition has a contribution of 80% towards critical thinking skills. As for mathematical reasoning component, it contributes 75% towards critical thinking skills, and the interpretation contributes 55% towards critical thinking skills. Then the component of analysis is 60%, the component of evaluation is 55%, and the component of inference is 48%.

Suggestions

Based on the analysis, it is suggested that (1) the Board of Educational Assurance conducts training on the formulation of instruments of critical thinking for teachers; (2) teachers implement formative evaluation of critical thinking in mathematics teaching–learning process in junior high school; (3) further researches by using data analysis of 2PL or 3PL are performed.

Reference

Adams, R.J., & Khoo, S. (1996). *Quest: The interactive test analysis system*. Melbourne: Australian Council for Educational Research.

Basilevsky, A. (1983). *Applied matrix algebra in the statistical sciences*. Amsterdam: Elsevier Science Publishing

Ben-Zvi, D., & Garfield, J. (2005). *The challenge of developing statistical literacy, reasoning and thinking*. New York, NY: Kluwer Academic Publisher.

Bjorklund, D.F. (2005). *Children's thinking*. Belmort, CA: Wadsworth Thomson Learning.

Cheng, S., Liu, E., Lin, C., Teng, S., Tung, Y. (2008). Learning design on networked peer assessment for junior high. *Proceeding of the 7th WSEAS International Conference on e-Activities*. ISSN: 1790-5117.

Crowley, M.L. (1987). The van Hiele model of the development of geometric thought. In M. M. Lindquist and A. P. Shulte (eds.). *Learning and teaching geometry, K-12, 1987 Yearbook*. Reston, VA: National Council of Teachers of Mathematics, 1–16.

Ennis, R. (1985). A logical basis for measuring critical thinking skills. *Educational Leadership, 43*(2), 44-48.

Ennis, R. (1993). Critical thinking assessment. *Theory into Practice*, 32(3), 179-186.

Ennis, R. (2011). Critical thinking: Reflection and perspective part I. *Inquiry Critical Thinking across the Discipline*, 26(1), 4-18.

Ennis, R., & Weir, E. (1985). *The Ennis-Weir critical thinking essay test*. Pacific Grove, CA: Midwest Publications.

Facione, P.A., Giancarlo, C.A., & Gainen, J. (1995). The disposition toward critical thinking. *Journal of General Education*, 44(1), 1-25.

Facione, P.A. (2000). The disposition toward critical thinking: Its character, measurement, and relationship to critical thinking. *Informal Logic*, 20(1), 61-84.

Facione, P. (1990). *California critical thinking skills test: College level*. Millbrae, CA: California Academic Press

Falchikov, N. (2005). *Improving assessment through student involvement: Practical solutions for aiding learning in higher and further education*. New York, NY: Routledge. ISBN-0-415-30821-6.

Ferdinand, A. (2002). *Structure equation modeling dalam penelitian manajemen* [Structure equation modeling in management research]. Semarang: Badan Penerbit Universitas Diponegoro.

Fischer, K.W. (1980). A theory of cognitive development: The control and construction of hierarchies of skills. *Psychological Review*, 87(6).

Gall, M.D., Gall, J.P., & Walter, B.R. (2003). *Educational research: An introduction* (7th ed.). Boston: Allyn and Bacon.

Ghozali, I. & Fuad. (2008). *Structural equation modeling, teori, konsep, dan aplikasi dengan program Lisrel 8,80* [Structural equation modeling theory, concept, and application with Lisrel 8.80 program]. Semarang: Badan Penerbit Universitas Dipenogoro.

Hair, J.F., Black, W., Babin, B.J., & Anderson, R.E. (2010). *Multivariate data analysis (7th ed.)*. Upper Saddle River, NJ: Pearson-Prentice Hall.

Halpern, D.F. (2010). *Halpen critical thinking assessment*. Amsterdam: Schuhfried Gmbh.

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

Kawashima, N. & Shiomi, K. (2007). Factors of thinking disposition of Japanese high school students. *Social Behavior and Personality*, 35(2), 187-194.

Kartowagiran, B. (2008). Validasi dimensionalitas perangkat tes ujian akhir nasional SMP mata pelajaran

Matematika 2003-2006. *Jurnal Penelitian dan Evaluasi Pendidikan*, 12(2), 177-195. Retrieved from http://journal.uny.ac.id/index.php/jpep/article/view/1426.

Kieran, C. (2004). Algebraic thinking in the early grades: What is it?. *The Mathematics Educator,* 8(1), 139-151.

Kline, T.J.B., & Klammer, J.D. (2001). Path model analyzed with ordinary least square multiple regressions versus Lisrel. *The Journal of Psychology,* 135(2), 213-225.

Mardapi, D. (2008). *Teknik penyusunan instrumen tes dan nontes* [Technique of test and non-test instrument arrange-ment]. Yogyakarta: Mitra Cendikia.

Mason, M. (2007). Critical thinking and learning. *Educational Philosophy and Theory*, 39(4), 339-349.

Nieto, A.M., & Saiz, C. (2011). Skills and disposition of critical thinking: Are they sufficient?. *Analesde Psicologia*, vol 27(1), 202-209.

Plomp, T. (2007). Educational design research: An introduction to educational design research. *Proceedings of the seminar conducted at the East China Normal University, Shanghai (PR China)*, November 23-27.

Siegler, R.S., & Alibali, M.W. (2005). *Children thinking (4th ed.)*. Upper Saddle River, NJ: Pearson Prentice Hall.

Siegel, H. (2010). Critical thinking. *International Encyclopedia of Education*, 6, 1441-1445. Auto's personal copy. This article draws extensively upon Siegel 2003 and Bailin and Siegel 2003.

Schumacher, R.E., & Lomax, R.G. (1998). *A beginner's guide to structural equation modeling* (2rd ed.). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Thompson, B. & Vacha-Haase, T. (2000). Psychometrics is datametrics: The test is not reliable. *Educational and Psychological Measurement*. 60, 174-195.

Van der Linden, W.J. & Hambleton, R.K. (1997). *Handbook of modern item response theory*. New York, NY: Springer-Verlag New York, Inc

Van de Walle, J., Karp, K., Bay-William, J. (2011). *Elementary and middle school mathematics: Teaching develops mentally*. Boston, MA: Allyn & Bacon.

Wijanto, S.H. (2007). *Structural equation modeling dengan Lisrel 8.8* [Structural equation modeling with Lisrel 8.8]. Jakarta: Graha Ilmu.