

## Development of CIPP Model of Evaluation Instrument on the Implementation of Project Assessment in Science Learning

Jati Aurum Asfaroh<sup>a</sup>, Dadan Rosana<sup>a</sup>, Supahar<sup>a</sup>

<sup>a</sup>Yogyakarta State University, Yogyakarta, INDONESIA

### ABSTRACT

The aim of this study to develop an instrument of evaluation constructed by CIPP model on the implementation of performance assessment in science learning. This study used research and development (R & D) method; adapting 4D Model modified by the development of non test instrument. The subject of this developing study were 8<sup>th</sup> grade students and science teachers of Junior High School in Yogyakarta. Validity test is done by the content, construct, and language; its analyzed by using V'Aiken formula. Reliability testing is done by Inter Class Correlation (ICC) technique. The validity results indicates that the instrument as a valid. The V'Aiken coefficient ranged from 0,86 to 1,00. The alpha reliability coefficient for observation sheet is 0,923 and questioner is 0,916, its categorized reliable. The result shows that an evaluation instrument constructed by CIPP model categorized not only as valid in content, construction, and language, but also reliable. The instrument of evaluation used to evaluate the implementation of project assessment and the result show that the implementation of project assessment in Junior High School in Yogyakarta categorized as a good.

### KEYWORDS

Evaluation instrument, CIPP model, Project Assessment.

### ARTICLE HISTORY

Received 20 September 2017

Revised 01 October 2017

Accepted 10 October 2017

### Introduction

The nature and extent of students, understanding of scientific concepts and phenomena are key components of any science curriculum (Treagust, 2006). Science is defined as the knowledge gained through data collection by experimentation, observation and deduction to produce an explanation of a phenomenon that can be trusted. Learning is an effort to learners with the aim of improving the knowledge of learners to be more independent in the face of

**CORRESPONDENCE** J.A. Asfaroh ✉ [jati.aurum2015@student.uny.ac.id](mailto:jati.aurum2015@student.uny.ac.id)

© 2017 J.A. Asfaroh, D. Rosana & Supahar

Open Access terms of the Creative Commons Attribution 4.0 International License apply. The license permits unrestricted use, distribution, and reproduction in any medium, on the condition that users give exact credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if they made any changes. (<http://creativecommons.org/licenses/by/4.0/>)

global developments. In accordance with the established National Education Standards Board, that Natural Science is concerned with how to systematically find out about nature so that science is not only the mastery of a collection of knowledge in the form of facts, concepts or principles, but also is a process of discovery to gain a deeper understanding. From the explanation, it can be said that the learning activities are required to be oriented to the learner activity in the classroom and learning is focused on the process, not on the results achieved. Learners are required to build their knowledge through the concept discovery process with the teacher only as a facilitator.

In line with the nature of science that current learning activities of science dispected oriented to the learner activities. Natural science learning should be able to generate curiosity or curiosity of learners about natural phenomena studied, so that this kind of learning activity can make learners become accustomed in solving problems encountered. With so much curiosity of a learner against the natural phenomena that have been observed, of course, required a learning process in which learners can find their own answers to the natural phenomenon. Problem-solving skills is one of the high-level skills that must be mastered by learners to be able to keep up with the times. As expressed by Wismath et al (2014: 1) that problem solving is an important component of 21st century learning. Some 21st century global problems are in fact closely related to the provision of learning that is beneficial to the future of learners. The problem-solving process is suitable to be applied in science learning because it can improve students' logical, critical, creative and innovative thinking skills. It is seen that learners can learn to solve a problem of its own after the implementation of the project activities in Dewey's schools (Akinoglu, 2008: 2). Further explanation from Priyambodo and Wiyarsi (2011), project assessment is a thorough assessment of the learners' ability through the task containing the investigation and should be resolved within a specified time. This assessment project was to assess the general skills of investigating, understanding and knowledge in a particular field, the ability to apply knowledge in an investigation and to assess the ability to inform the subject clearly. In general, learning was designed with the application of project-based assessment which provides more opportunity on learners to actualize themselves during learning. In a scientific approach, learning assessment process applies authentic assessment, in the form of project assessment

Interviews that have been conducted with science teachers SMP N 1 Galur and SMP N 1 Sleman, assignment of the project is often done in school, but problem-solving skills can not be measured and well documented. Teachers have used project assessment, but no evaluation of the project assessment instrument has been used. These specific skills require assessments that have specific competence criteria so that the learning objectives can be achieved by the students.

One of the most important components of learning is the availability of learning assessment instruments to measure the success of learning that has been done. To measure the achievement of the IPA learning process, an appropriate assessment instrument is needed. At this stage of the learning process, assessment is an important aspect and needs to be considered. Assessment (assesment) is an activity to make measurements and benchmarking the learning outcomes of learners. Assessment activities will make it easier for teachers to determine appropriate follow-up, so a valid

instrument is required with appropriate and clear criteria to produce an accurate assessment of the learner's abilities. Assessment of learning outcomes of learners can be the result of learning knowledge (cognitive), skills, and attitudes. Assessment activities can be done with test and non-test techniques, while the instrument used can be either a question or non-question. "Assessment is one continuous process to determine method proposed applies in lessons by teachers and at the same time supervisors can give guidance and guide to them to overcome the shortage that exist" (Yahaya, 2001).

There is an assessment standard in Curriculum 2013 that stipulated in Permendikbud No. 23 of 2016 which says "Assessment is the process of collecting and processing information to measure the achievement of learning outcomes of students. Objective assessment of learning outcomes by educators aims to monitor and evaluate the process, the learning progress, and improvement of learning outcomes of students on an ongoing basis. Assessment of learning outcomes of students in primary education and secondary education includes aspects, attitudes, knowledge and skills. Skills assessment is an activity undertaken to measure learners' ability to apply knowledge in performing certain tasks. Skills assessment done through practice, product, project, portfolio, and / or other techniques in accordance with the competencies assessed ". It can be said that educators must be able to assess the student's ability and learning outcomes not only on the cognitive aspects but also to measure the attitude and skills.

According to Sterling (2005: 33), effective assessment should be related to the way of learning, and the results can be used to inform the learning outcomes. Curriculum 2013 stresses on authentic assessment. In Permendikbud No. 66 of 2013, it made clear that authentic assessment is a comprehensive assessment conducted to assess an input, process and output in learning. Based on the statement, authentic assessment is a meaningful measurement is significantly above the learning outcomes of students. Besides authentic assessment is expected to involve students actively in the learning process because students were asked to reflect and evaluate their own performance in order to increase a more mature understanding of the purpose of learning and encourages higher learning ability. In an authentic assessment, teachers can appreciate the skills, attitudes and knowledge of what is and is not owned by learners, how they apply their knowledge in daily life and so on. According to (Callison, 1998), it provides broader understanding of the authentic assessment, namely: Authentic assessment is an evaluation process that involves multiple forms of performance measurement reflecting the student's learning, achievement, motivation, and attitudes on instructionally-relevant activities. Examples of authentic assessment techniques include performance assessment, portfolios, and self-assessment. From the above statement, it can be summarized that authentic assessment is an evaluation process to measure performance, achievement, motivation, and attitudes of students in relevant learning activities. Examples of assessment applied as performance assessment, portfolio and self-assessment. Authentic assessment is defined as an assessment of the product and performance in relation to real-life experiences of students.

Project assessment is an appropriate and valid assessment tool for measuring and assessing student problem solving skills in schools. The development of the project assessment instrument has been carried out by the first researcher Sukmasari (2016) by developing a project assessment

instrument to measure problem-solving skills. The results obtained are the project assessment instrument to measure the problem solving capability stated valid and reliable, but the instrument has not been tested its characteristics when the implementation is widely and has not been used more widely by teachers of science SMP. Limited project assessments have not been able to demonstrate standardized instrument criteria, so it is necessary to disseminate their use. A valid and reliable valuation instrument should be an advantage for the teacher to use the instrument so that the assessment activity gets better. The project assessment instrument has a weakness in the case of implementation because it often does not conform to the characteristics of the material so that an evaluation is required before it is disseminated.

Assessment instruments in the form of a broad assessment of the project as a form of efforts to disseminate products that have been declared valid and reliable, it is necessary to conduct evaluation activities in its application. Evaluation of the implementation of the assessment program is intended to determine the achievement and effectiveness of the use of the project assessment instrument to measure problem-solving skills. The results of this program evaluation will be expected to provide a decision that can be used to follow up the application of assessment instruments in schools in a broad scope. Evaluation of learning is a process of assessment to take a decision which is based on a comprehensive assessment of measurement results include: affective (attitude), cognitive (knowledge), and psychomotor (skills). For the evaluation instruments should be able to provide the measurement and comprehensive assessment, covering all aspects of the learning outcomes of students.

Evaluation activities require evaluation tools / instruments to evaluate a program. Appropriate evaluation instruments will produce evaluation results that are appropriate to the purpose of the evaluation. Appropriate evaluation instruments required needs analysis. The program consists of at least three components, namely input, process and output. The CIPP evaluation model stands for context, input, process and product. The CIPP evaluation model consists of four components: contexts, inputs, processes and outputs (Phattharayuttawat, 2009).

This research uses CIPP evaluation model (context, input, process, product). The components of the contexts in this study include the appropriateness of the project assessment instrument with KI & KD. Input components include the understanding and skills of teachers using project assessment instruments as well as learning environments. The process components include the implementation of the project assessment instrument. The output component includes an analysis of the results of the implementation of the project assessment, so that with the results of the information can be determined follow-up actions of the project assessment in the future. The analysis of the needs of the evaluator will gain clarity of the issues in the assessment so as to provide recommendations to policy makers.

The evaluation instrument used should have good instrument characteristics. The use of evaluation instruments in accordance with the evaluation characteristics to be performed will result in valid information, so that this information can be used to make decisions regarding the follow-up to the implementation of project assessment. The developed instrument must follow the standardization step of an evaluation instrument (Arifin, 2013). The

CIPP model is chosen because the model is effective for obtaining formative and summative evaluation results and for obtaining decisions and problem solving.

The CIPP evaluation model is systematically designed as an evaluation guide in compiling relevant questions and early behavioral assessments of the project (context and input evaluation), while on project implementation (input and process evaluation) and at the end of the project (product evaluation) (Guili, 2011). A survey by the American Society for Training Development shows that the CIPP model is preferred over other evaluation models. Conditions in real life have not been made the development of CIPP model evaluation tools project assessment of the application of project assessment on science learning to measure problem solving skills.

In conducting the evaluation, it needs tools / instruments to evaluate a program. Proper evaluation instruments will create evaluation results in accordance with the purpose of the evaluation. To produce a proper evaluation instruments, it would require a needs analysis. A minimum program consists of three components, namely input, process and output. An understanding of the instrument to be important for the evaluation and assessment practices; in general, teachers always applied the measurement process. Good instrument is an instrument that meets the requirements or specific rules, to provide accurate meaningful data according to function, and the only measure samples of certain behaviors. According to (Arifin, 2013), the characteristics of a good instrument are valid, reliable, relevant, representative, practical, descriptive, specific, and proportionate. In line with the opinion of D. L. Stufflebeam (1984: 3) defines the evaluation is defined as a process to describe, obtain and provide information that is useful to assess alternative decision. Terry D. TenBrink (1974: 8) Evaluation is the process of obtaining information and using it to form judgments which in turn are to be used in decision making. Furthermore, according to Arifin (2013), he found an ongoing process sustainable and systematic way to determine a quality that includes the value and meaning of something. This quality determination based on certain criteria and considerations.

Evaluation instrument development research on the application of this Science teaching project appraisal used CIPP evaluation model. This evaluation model is considered suitable to apply. "CIPP assessment's models selected because his effectiveness to get revenue formative and summative and to find decision and problem solving ability" (Yahaya, 2001). It can be summarized that the evaluation model of CIPP was chosen because it is effective to obtain the results of formative and summative as well as to determine the decision and prowess problem solving. According to Guili (2011) "in education setting, the CIPP evaluation model has been used to evaluate numerous educational projects and entities". This is supported by Safruddin (2014) which states that the CIPP evaluation model considers a program as a system, so that the evaluation of the program as the system is to be executed in detail based components.

The components of the system will be evaluated by the CIPP model of evaluation instruments are as follows: Context, Stufflebeam (1984) in his book entitled *Systematic Evaluation*. Meanwhile, according to Oliva (1992), an evaluator establish a situation where the system will be evaluated and perform an analysis of unmet needs, as well as identify any reason why the background for those requirements have not been or cannot be reached. Evaluation in the

context section focuses on the evaluation of activities related to the needs analysis, needs that have been achieved or not achieved. The analysis also examined in more detail that is by looking for reasons for the fulfillment of those needs. In addition to determining the needs of the program, in this context evaluation may also be specified program objectives.

These components include determining the objectives, missions and objectives of the school's evaluation (Patil & Kalekar, 2015). In this evaluation instrument development research, context component conformity assessment instrument includes projects that have been developed in the first year using KI and KD.

*Input*, Stufflebeam (1984) "... the input evaluation structuring decision ...". It can be summarized that the input evaluation is the framework of the decision. The evaluation aims to help regulating a decision, determine the sources, an alternative which will be taken, what plans and strategies to achieve the needs, and how the work procedures to achieve. The same is expressed by Stufflebeam (1984) which states that this evaluation helps to determine the information that will be used to meet the objectives or needs. Crunkilton (1979) argued that the input evaluation is closely related to the source determination and any strategy that will be used to achieve the purpose of the system or program being evaluated. Examples of sources that affect efforts to achieve the goal are how teachers teach, the use of instructional media and learning environment.

Based on the description above, it can be seen that the input evaluation related to what strategies can be used to achieve the needs that have not been or cannot be reached. The strategy may come from teachers teaching through teaching skills and the use of media to learn, and come from students in the form of student' spirit, concentration, and understanding.

*Process*, According to Oliva (1992) states that the evaluation process with regard to the reciprocal arising from the application of a system or program being evaluated. Furthermore Stufflebeam (1984) revealed the "process evaluation, to serve implementation decision" process of evaluation activity aims to help implementing the decision. Furthermore Crunkilton (1979) revealed that the evaluation process is very closely associated with learning and evaluation process is focused on how effectively the effects of the application of the system or program being evaluated. The description above describes the evaluation process is based on several experts, so it can be seen that the evaluation process is an evaluation of activities which focus on how a system or program being evaluated.

*Product*, evaluation on results or product has the purpose to know the results of what has been achieved from the implementation of the system or program, and follow-up what will be done after the system or specific programs implemented (Arifin, 2013). Crunkilton (1979) states that information obtained from the evaluation of this product is very important because the evaluation results will determine the follow-up to be taken. Furthermore Stufflebeam (1984) "product evaluation, to serve recycling decision". In this component, there should be the result has been achieved from a program and what to do after the program runs. The evaluation is aimed to help the next decision.

Product evaluation related to the analysis toward results of the implementation of the system or program. The information obtained can show the results of what has been achieved so that information can be determined by

the next steps taken to follow up systems or programs in the future. The CIPP evaluation model seen from the learning concept has the quality of learning process in Science Technology and Society (STM) by influenced input, process and product (Issac, 1971). Stufflebeam's evaluation model is recommended to be a systematic framework as a reference to the conceptions, design, implementation and assessment of a program and to provide feedback and assessment of the effectiveness of the project for sustainable improvement.

Based on the above explanation, the researcher is interested in conducting research related to the development of CIPP model evaluation model in applying the assessment of science learning project to measure the problem solving skills of junior high school students more broadly, that is in Yogyakarta city area. The development of CIPP model evaluation instruments on broader project assessment for the city area of Yogyakarta, is expected to provide good results in the form of knowledge about the quality of instruments and the goodness of good project assessment instruments so as to improve the quality of science learning assessment in junior high. In this discussion, it will be limited to the validation and reliability of CIPP model evaluation instruments on the assessment of assessment of science learning projects to measure problem-solving skills.

### **Purpose**

The purpose of this study is to develop CIPP evaluation instrument on the implementation of performance assessment in science learning that tested its validity and reliability. In this study limited discussion on the validation and reliability of the CIPP model evaluation instrument on the implementation of project assessment to measure problem solving skills.

### **Method**

This research used 4D development model (Thiagarajan, 1975: 5) which is modified with nontes instrument development model (Rusilowati, 2013: 7-11). Development procedures include; 1) the define stage begins with a preliminary study of interviews; 2) the design phase includes designing the CIPP model evaluation instrument; 3) the development stage includes the development of non-test instruments namely instrument review and assembling instruments; 4) disseminate stage in this research is done by delivering the product to the Principal. (1) determine the scoring system, (5) analyze the instrument, (6) perform the test, (7) analyze the instrument, (8) assemble the instrument, (9) carry out the measurement, and (10) interpret the measurement result.

### **Result and Discussions**

#### ***Validity of CIPP Model Evaluation Instruments***

Product validation is done by experts and practitioners involving two lecturers, two science teachers of SMP class VIII, and three colleagues of graduate students of science education. Validation is performed to determine the validation of the contents of CIPP model evaluation instruments, including the compatibility between statements with indicators and communicative language or not. Validation by these experts and practitioners refers to content validation. This validation process uses the validation sheet that each validator must fill in. The validation sheet component used by the validator to review the contents of

the CIPP model evaluation instrument includes aspects of substance, language aspects, and construction aspects.

Data obtained from this content validation stage is the valuation and input from the validator. The validator is required to provide a check mark for each statement in the CIPP model evaluation instrument in accordance with the statement contained in the validation sheet. The check mark provided by the validator has a score of 1. The overall score given by the validator, both for the substance, construction, and language aspects is then analyzed using Aiken's formula. This Aiken's formula is used to determine the content validity coefficient ( $V$ ) in each item of statement in the CIPP model evaluation instrument. The results obtained from this validator are then converted into four categories, valid categories without revisions, valid with little revisions, valid with many revisions, and invalid (Lynn, 1986). In addition to the assessment of the scores analyzed with Aiken's formula, the validator also provided input on the CIPP model evaluation instrument that can be used to revise and improve the quality of the CIPP model evaluation instrument being developed.

The CIPP model evaluation instrument consists of an observation sheet and a questionnaire. Observation sheets in the instrument include the observation sheets for the input, and process aspects, while the questionnaire in the evaluation instrument is a questionnaire for aspects of context, process and product. The scoring analysis of the validators using Aiken's formulas for observation sheets and questionnaires in the evaluation instrument was confirmed by the number limit on the Aiken's  $V$  table for the number of categories of span four and number of rater eight, ie 0.78 (1985: 134). Rater involved in the validation process of evaluation instrument contents are two lecturers of evaluation and material experts, two science teachers of SMP class VIII, and three graduate students of IPA. The validation results of the CIPP model evaluation instrument for each item of statement on the observation sheet on the context, input, process, and product aspects are presented in Table 1.

**Table 1.** Validation Results of Evaluation of CIPP Model Evaluation Instrument

Aspect	Statement Number	V Aikens	Category
Context	1,3,5,6	0,95	Valid
	2,4,7	1,00	
Input	1a, 1b, 1c, 1d, 2a, 2b, 2c, 2d, 3b, 3c, 3d, 4a, 4b, 4c, 5a, 5c, 6a, 6b, 6c, 6d, 7a, 7b, 7c, 8a, 8b, 8c, 8d, 9b, 9c, 9d, 10b, 10c, 10d	1,00	Valid
	4d, 5d, 9a,10a	0,90	
	3a, 5b, 7d	0,86	
	1c, 1d, 2b,2c,2d, 3b, 3c, 3d,	1,00	
	1b, 2d	0,90	Valid
	1a, 2a, 3a, 4a,	0,86	
Process	1a, 1b, 1c, 1d, 2a, 2b, 2c, 2d, 3a, 3b, 3c, 3d, 4a, 4b, 4d, 5a, 5b, 5c, 5d, 6b, 6c, 7c, 7d, 8a, 8b, 8c, 8d, 9b, 9c, 9d, 9a, 9b, 9c	1,00	Valid
	4c, 7b, 9a	0,95	
	6a, 6d, 7a, 10d	0,86	
	2, 3, 4, 6, 7, 8, 10, 11, 12, 14, 15, 16, 18, 20, 22, 23, 24	1,00	
	19	0,95	Valid
	9, 17, 21, 25	0,90	
	1, 5, 13,	0,86	
Product	1, 3, 4, 6	1,00	Valid
	5, 9	0,90	
	2, 7, 8	0,86	

Experienced content validation results in the form of content validity stated in V aiken and produced some suggestions and inputs by experts related to the improved quality of evaluation instruments of the developed CIPP model.

The expert advises against the preparation of sentences in the statement item so that the statement statement is effective.

The result of instrument validation analysis shows that the evaluation instrument of CIPP model is included in valid category with V'Aiken coefficient 0,86-1,00. A valid instrument affects the data or information collected in the field. This is supported by the statement of Darmadi Hamid (2011: 115) which states that the use of good instruments by researchers able to collect data or information from the object or subject studied.

### **Reliability of CIPP Model Evaluationa Instruments**

Reliability is one of the requirements of the instrument can be said to be feasible and can be used for measurement on field trials. Reliability can be analyzed by ICC (interclass correlation coefficient) technique. The analysis was carried out by involving the scores of the evaluation results of the appraisal of the five evaluators. The evaluation score of the application of project appraisal in SMPN 1 Galur class VIII B was analyzed by ICC technique. Reliability of evaluation instruments can be known from the value of Cronbach Alpha. Based on analysis with ICC technique, Cronbach Alpha coefficient for context aspect questionnaire is 0,916, while Cronbach alpha coefficient for observation sheet is 0,923. The Cronbach Alpha coefficient shows that the reliability of the evaluation instrument consisting of observation sheets and questionnaires, including in the category is very good. The category of reliability of this instrument in accordance with the opinion of Gliem & Gliem (2003: 87) which divides the Cronbach Alpha coefficient category into 6, namely  $\geq 0.9$  (Excellent),  $\geq$  (Good)  $\geq 0.7$ ,  $\geq 0.6$  (Questionable)  $\geq 0.5$  (Poor), and  $<0.5$  (Unacceptable). This valid and reliable instrument of evaluation is valid for further measurement. The results of the reliability analysis of CIPP model evaluation instruments are presented in Table 2.

**Table 2.** Results of Reliability Analysis of CIPP Model Evaluation Instruments

<b>Instruments</b>	<b>Aspect</b>	<b>Cronbach's Alpha Value</b>	<b>Category</b>
Observation sheet	<i>Input,</i>	0,923	Reliabel
Questionnaire sheet	<i>Context,</i>	0,916	

Revisions in this study were conducted twice, based on (1) input and suggestion from the validator and (2) shortages during the trial. Suggestions and feedback are used to improve the CIPP model evaluation instrument product to make it feasible to use. Revision of the CIPP evaluation model instrument: (a) improvements in writing a few words do not conform to Indonesian regulations; (b) improvements in language selection are less effective; (c) the improvement of the statement item on the CIPP model evaluation instrument on several indicators; (d) the improvement of a statement item that is less concise, clear, and assertive; (e) improvement in unnecessary

items of statements; (f) the improvement of the assessment rubric is less clear, and (g) the refinement of the item is still looped.

### Conclusion

Based on the study results and discussion, it can be concluded. The CIPP model evaluation instrument has several characteristics, including: CIPP model evaluation instruments including valid categories, and reliable. The validity of instrument contents based on the experts judgments is shown by the coefficient of V Aiken with the range 0.86 to 1.00. The reliability of the CIPP model evaluation instrument is shown by the Cronbach Alpha coefficient of 0.916 for the questionnaire and 0.923 for the observation sheet. Based on these results, CIPP model evaluation instruments fall into the category reliabel

### Disclosure statement

No potential conflict of interest was reported by the authors.

### Notes on contributors

Yogyakarta State University, Yogyakarta, Indonesia

**Jati Aurum Asfaroh**, Universitas Negeri Yogyakarta.

**Dadan Rosana**, Universitas Negeri Yogyakarta.

**Supahar**, Universitas Negeri Yogyakarta.

### References

- Aiken, L. R. (1985). Three Coefficients for Analyzing The Reliability and Validity of Ratings. *Educational and Psychological Measurement*, 45, 131 dan 134
- Akinoglu, O. (2008). *Assessment of the inquiry-based project implemetation process i science education student point of view. Intrnational Journal Of Instruction*, 1-12.
- Arifin, Z. (2013). *Evaluasi Pembelajaran*. Bandung: PT Remaja Rosdakarya.
- BSNP. (2010). *Laporan BSNP: Pengembangan paradigma pendidikan nasional abad XXI*.
- Callison, d. (1998). *Authentic Assessment. Scholl Library Media Activities Monthly 14 no 5*. Dipetik Maret 01, 2017, dari <http://www.ala.org/aasl/sites/ala.org/aasl/files/content/aaslpubsandjournals/slr/edchoice/SLM Q AuthenticAssessment InfoPower.pdf>.
- Crunkilton, C. (1978). *Curriculum development in Vocational and Technical Education*. Boston: Allyn and Bacon Inc.
- Gliem, J. A & Gliem, R. R. (2003). Calculating, Intepreting, and Reporting Cronbach's Alpha Reliability Coefficient for Likert-Type Scales. *Midwest Research to Practice Conference in Adult, Continuing, and Community Education*.
- Guili, N. Z. (2011). *Using The Context, Imput, process and Product Evaluation Model (CIPP) as a Comprehensive Framework to Guide the Planning, aimplemetation, and assessment of Service-learning Programs. Journal of Higher Educational outreach and Engangement, Volume 15 Number 4*, 57.
- Hamid, D. (2011). *Metode Penelitian Pendidikan*. Bandung: Alfabeta.
- Lynn, M. R. (1986). Determination and Quantification of Content Validity. *Nursing Research*, -, 382.
- Oliva, P. (1992). *Developing The Curriculum Third Edition*. New York: Harper Collins. .
- Patil, Y., & Kalekar, S. (2015). *Cipp model for school evaluation*. June-July, 2(10), 2615–2619. Retrieved from [www.srjis.com](http://www.srjis.com)

- Phattharayuttawat, S. (2009). *An evaluation of the curriculum of a graduate programme in Clinical Psychology*. South East Asian Journal of Medical Education, 3(1), 14–19. Retrieved from [http://seajme.md.chula.ac.th/articleVol3No1/OR2\\_Sucheera Phattharayuttawat.pdf](http://seajme.md.chula.ac.th/articleVol3No1/OR2_Sucheera%20Phattharayuttawat.pdf)
- Republik Indonesia. (2013). *Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia Nomor 66 Tahun 2013 tentang Standar Pendidikan*
- Republik Indonesia. (2016). *Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia Nomor 23 Tahun 2016 tentang Standar Pendidikan*.
- Rusilowati, Ani. (2013). *Pengembangan Instrumen Nontes*. SNEP Tahun 2013 ISBN 978-602-14215-0-5. Diakses pada tanggal 14 Oktober 2016, dari <http://conf.unnes.ac.id/index.php/snep/I/paper/viewFile/7/2>.
- Safruddin, S. A. (2014). *Evaluasi Program Pendidikan*. Jakarta: PT Bumi Aksara.
- Sterling, D. (2005). *Assessing Understanding*. *Journal Science Scope*, Vol. 28 N4 p.33-37.
- Stufflebeam, D. (1985). *Systematic Evaluation "A Self-Instruction Guide to Theory and Practice"*. Hingham Massachusetts: Kluwer Academic Publishers.
- Sukmasari, V. P. (2016). *Pengembangan Instrumen Project Assesment Untuk Mnegukur Keterampilan Pemecahan Masalah*. Yogyakarta.
- TenBrink, T. D. (1974). *Evaluation A Practical Guide for Teacher*. Columbia: McGraw-Hill.
- Thiagarajan, Sivasailam, Semmel, Dorothy S., Semmel, Melvyn I. (1975). *Instructional Development for Training Teachers of Exceptional Children*. Bloomington, Indiana: Indiana University.
- Treagust, David F. (2006). *Diagnostic assessment in science as a means to improving teaching, learning and retention 2006 Australian Conferece on Science and Mathematics Education* Sydney. UniServe Science Conference. Sydney.
- Wismath, S., Orr, D., & Zhong, M. (2014). *Student perception of problem solving skills*. *Teaching & Learning Journal*, 7(3),1-17.
- Wiyarsi, A. (2011). *Penilaian Proyek sebagai Implementasi Authentic Assessment untuk Meningkatkan Kemampuan Berpikir dan Kerja Ilmiah mahasiswa*. Dari [http://staff.uny.ac.id/sites/default/files/132312678/makalah%20 authentic%20assessment.pdf](http://staff.uny.ac.id/sites/default/files/132312678/makalah%20authentic%20assessment.pdf) diakses pada tanggal 20 Oktober 2016.
- Yahaya, A. H. (2001). *The using of Model Context, Input, Process and Products (CIPP in Learning Programs Assessment)*. *Inrenational Conference on Challenges and Prospects in Teacher Education* (hal. 1). Johor: Universiti Teknologi Malaysia.