Development of the evaluation instrument use CIPP on the implementation of project assessment topic optik

Jati Aurum Asfaroh, Dadan Rosana, and Supahar

Citation: AIP Conference Proceedings **1868**, 080006 (2017);
View online: https://doi.org/10.1063/1.4995190
View Table of Contents: http://aip.scitation.org/toc/apc/1868/1
Published by the American Institute of Physics

Development of The Evaluation Instrument Use CIPP on The Implementation of Project Assessment Topic Optik

Jati Aurum Asfaroh [1, a)] , Dadan Rosana [2, b)] , Supahar [2, c)]

[1]*Science Education Study Program of Graduate School, Yogyakarta State University, Indonesia*
[2]*Faculty of Mathemathic and Natural Science, Yogyakarta State University, Indonesia*

[a)] Corresponding author: jatiaurumasfaroh@gmail.com
[b)] danrosana@uny.ac.id
[c)] pahar.fis@gmail.com

Abstract. This research aims to develop an evaluation instrument models CIPP valid and reliable as well as determine the feasibility and practicality of an evaluation instrument models CIPP. An evaluation instrument models CIPP to evaluate the implementation of the project assessment topic optik to measure problem-solving skills of junior high school class VIII in the Yogyakarta region. This research is a model of development that uses 4-D. Subject of product trials are students in class VIII SMP N 1 Galur and SMP N 1 Sleman. Data collection techniques in this research using non-test techniques include interviews, questionnaires and observations. Validity in this research was analyzed using V'Aikens. Reliability analyzed using ICC. This research uses 7 raters are derived from two lecturers expert (expert judgment), two practitioners (science teacher) and three colleagues. The results of this research is the evaluation's instrument model of CIPP is used to evaluate the implementation of the implementation of the project assessment instruments. The validity result of evaluation instrument have V'Aikens values between 0.86 to 1, which means a valid and 0.836 reliability values into categories so well that it has been worth used as an evaluation instrument.

INTRODUCTION

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. In line with this definition, "the nature and extent of students' understanding of scientific concepts and phenomena are key components of any science curriculum" [1]. Learning science is the process of interaction between learners and teachers, learning resources, media learning, and the learning environment to achieve the learning competencies of science that has been set [2]. The learning process includes several stages, namely planning, implementation process of learning, and assessment of learning outcomes. Science learning process is expected to cover three areas, namely knowledge, attitudes and skills. In science, it is important to assess the skills, because learning science requires an investigation, either by observation or experiment, as part of the scientific work that involves skills which is based on the scientific attitude. In connection with the Science objectives that the application of the concepts and skills of the Science process is needed to understand the nature and its contents as well as the students were able to solve the problems occurring in the surrounding environment. One of the skills that must be developed, measured and assessed in science is problem solving skills. Problem solving skills can be improved through the assignment of a project.

Assessment is an integral part of the learning process. The assessment includes the collection of information through a variety of assessment techniques and makes decisions based on the results of the assessment[3]. "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" [4]. It can be summarized that the assessment is a continuous process to determine the purpose of the methods used in teaching by teachers and at the same time supervisors can provide guidance and lead them to overcome existing shortcomings.

The 4th International Conference on Research, Implementation, and Education of Mathematics and Science (4th ICRIEMS)
AIP Conf. Proc. 1868, 080006-1–080006-9; doi: 10.1063/1.4995190
Published by AIP Publishing. 978-0-7354-1548-5/$30.00

There is an assessment standard in Curriculum 2013 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.

Effective assessment should be related to the way of learning, and the results can be used to inform the learning outcomes[5]. Curriculum 2013 stresses on authentic assessment. 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. 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[6]. 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.

One of assessment skills one is applied project assessment. This project would require the assignment of a specific instrument to measure students' thinking skills in the form of project assessment. Dewey explained that one of the most spectacular educational activities and effectively enables learning by doing and active learning in the inquiry-based science education is an activity of the project. The project plays an important role in enabling learners to understand the nature of Science and make the students enjoy learning Science. 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 [7]). Project assessment is a thorough assessment of the learners' ability through the task containing the investigation and should be resolved within a specified time[8]. 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.

Project assessment is an assessment instruments which is appropriate and valid to measure and assess the problem solving skills of students in the school. Evidenced by the first-year researchers, developing assessment project instruments is for measuring problem-solving skills[9]. The results obtained, project assessment to measure problem-solving ability is declared valid and reliable. However, these instruments have not been used more widely by the Junior High School Science teacher. The use of a limited assessment project has not been able to demonstrate the criteria of standardized instruments, so it should be the dissemination of its use. The existence of a valid and reliable assessment instrument should be able to be an advantage for teachers to use such instruments for better assessment activities.

The use of assessment instruments such as project assessment widely as an effort to disseminate the products that have been declared valid and reliable needs to be coupled with the holding of the evaluation activities in its application. Evaluation of the implementation of the assessment program is intended to determine the achievement and adherence to the use of project appraisal assessment instruments to measure problem solving skills. The results of the evaluation of this program are expected to give a decision which could be used to follow up the implementation of assessment instruments of assessment project at 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 can be used to measure the success of learning undertaken achievement indicators [10]. Indicators can be an indicator of cognitive learning products, cognitive processes, psychomotor, and affective [11]. Evaluation of learning is a process of assessment to take a decision based on the comprehensive measurement results include; affective (attitude), cognitive (knowledge), and psychomotor (skills). For the evaluation instruments should be able to provide the measurement and comprehensive assessment [12].

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. The characteristics of a good instrument are valid, reliable, relevant, representative, practical, descriptive, specific, and proportionate [13]. The evaluation is defined as a process to describe, obtain and provide information that is useful to assess alternative decision [14]. Evaluation is the process of obtaining information and using it to form judgments which in turn are to be used in decision maked[15]. Found an ongoing process sustainable and systematic way to determine a quality that includes the value and meaning of something[13]. This quality determination based on certain criteria and considerations.

Evaluation of an activity or program has several objectives relating to the enforceability of activities and programs. The purpose of evaluation is mentioned as below:

1. Provide information about the implementation of activities or programs that can be used as take decision.
2. Determine whether an activity or program succeed or fail, and can also be shown on the factors that affect a given environment.
3. Provide alternative solutions that can help improving an activity or program.
4. To provide an understanding and explanation of the characteristics of the activity or program and its implementation [16].

Evaluation aims to obtain information to determine the level of achievement of the objectives of the program through the knowledge of the enforceability of the program of activities [17]. In general, the evaluation aims to determine the level of success of an activity or program. The success rate of related knowledge is based on information obtained from the evaluation by using certain criteria.

Revealed that in evaluating the activities of the education program, there are several models of evaluation which are often used to evaluate education programs[17]. Some models of evaluation that is often used to evaluate educational programs, among others:

1. Goal Oriented Evaluation Model
2. Goal Free Evaluation Model
3. Formative Summative Evaluation Model
4. Countenance Evaluation Model
5. Responsive Evaluation Model
6. CSE-UCLA Evaluation Model
7. CIPP Evaluation Model
8. Discrepancy Model

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" [4]. 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 Zhang Guili "in education setting, the CIPP evaluation model has been used to evaluate numerous educational projects and entities"[18]. The theoretical framework is based on the CIPP Evaluation Model developed by Daniel Stufflebeam in 1970s [21]. This model was designed to evaluate programs, project, personel, products, institution or system from various disciplines such as education field. Housing and community development, transportation, safety and miltary personel review system[23]. This is supported 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[17]. The main reason why this evaluation model is chosen for this study is because this model is based on the managementoriented evaluation approach which helps the decision-makers to plan, implement and evaluate programs [19].Furthermore, it is widely used by the evaluators [20] and it covers wide variety of dimensions which could bechosen by the evaluators to best suit their studies. CIPP is an abbreviation for the evaluation: Context, Input,

Process, and Product. Context evaluation is used to choose the goal. Input evaluation is used to revise the plan. Process ealuation is used to guide the implementation of the plan. Product evaluation is used to provide the inspection determination[24][25][26].

The components of the system will be evaluated by the CIPP model of evaluation instruments are as follows:

1. Context

The primary orientation of a context evaluation is to identify the strengths and weaknesses of some object, such as an institution, a program, a target population, or a person, and to provide direction for improvement. A context evaluation also is aimed at examining whether existing goals[14]. Revealing that on this component, the main orientation of the evaluation context is to identify the strengths and weaknesses of the object, as an institution, a program, a target population or individual / person, and provide clues to progress. An evaluation of the context is also aimed on checking purposes. On this component, an evaluator identify several things, among others teacher factor, working facilities, student, work atmosphere, the role of the school committee, the community, as well as several other factors that have the possibility of influence on the system being evaluated[16]. The components of this context relates to the target population, the analysis of the needs of the institution, and the opportunity to realize the needs of the institution[13]. 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 [29]. Context evaluation to serve planning decision, which is the context of evaluation to help administrators plan decisions, determine program needs, and formulate program objectives[13].

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. 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.

2. Input

The input evaluation structuring decision[14]. 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 [13]. The same is expressed which states that this evaluation helps to determine the information that will be used to meet the objectives or needs[14]. 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 [27]. 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.

3. Process

The evaluation process with regard to the reciprocal arising from the application of a system or program being evaluated [22]. Process evaluation, to serve implementation decision" process of evaluation activity aims to help implementing the decision [14]. 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 [27].
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.

4. 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 [13]. Information obtained from the evaluation of this product is very important because the evaluation results will determine the follow-up to be taken [27]. 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 [14].

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.

Based on the explanation above, it is necessary to do research related to the development instrument CIPP model evaluation on the application of Science teaching project assessment to measure the problem-solving skills of students of SMP, more broadly, in the Yogyakarta area. Through the development of CIPP model evaluation

instrument in the broader project for the area of Yogyakarta, is expected to give good results in knowledge about the quality of assessment instruments and how good instrument projects can be done to improve the quality of Science teaching in Junior High School.

METHOD

The research type is Research and Development (R & D). Research and Development is a research model that is used to produce a specific product and test the effectiveness of the product. Products mentioned in this study is the CIPP model evaluation instrument on the implementation of projects in Science learning assessment to measure the problem-solving skills of Junior High School grade VIII students about optical of the human eye.

Development models used in this research is to integrate 4D model [28], with the non-test instrument development model [29]. Model development of 4D has steps: (1) define, (2) design, (3) develop, (4) disseminate, while the development model of instrument non-test has 10 steps as follows: (1) determine the specifications of the instrument, (2) write instruments, (3) determine the scale of the instrument, (4) determine the system of scoring, (5) identify instruments, (6) test, (7) analyze the instruments, (8) assemble instruments, (9) carry out measurements, (10) interpret measurement results.

The subject of the product of CIPP model evaluation instrument on the implementation of the project assessment is student of grade VIII SMP N 1 Galur Kulon Progo and SMP N 1 Sleman of second semester of academic year 2016/2017.

Analysis of the validation of the content of the descriptive and quantitative basis. Quantitative analysis using Aiken's V analysis [30] by the following formula:

$$V = \frac{\sum s}{n(c-1)} \tag{1}$$

Description :
s = $r - l$o
n = number of panels of assessors
lo = lowest validity assessment
c = highest validity assessment
r = the numbers given by an assessor

CIPP model evaluation instrument reliability in the application of science teaching project assessment on the topic Optics obtained using the estimation method of reliability Interclass Correlation Coefficients (ICC) SPSS 20.0 and the Cronbach's alpha coefficient (α) . Criteria of evaluation instruments mdel CIPP has good reliability if the ICC ≥ 0.60, and the coefficient α ≥ 0.70.

RESULTS AND DISCUSSION

Procedure Development of An Evaluation Instrument CIPP Model

Procedure development in this research is to integrate 4D model with the model of development of non-test instrument, which measures the non-test instrument development model sorted and combined into 4D development model. The explanation of procedure CIPP model evaluation instrument development on the application of science teaching project assessment to measure problem-solving skills as follows.

Preliminary Studies

Teacher competency analysis conducted to obtain information on the understanding of teachers and how to use the project by teacher assessment instruments. This analysis was conducted with interviews with some of the Junior High School science teacher. Analysis of material is done by studying KI and KD in Curriculum 2013. The competencies used in this study is KD 3.11 Describe the properties of light, shadow formation, as well as its application to explain human vision, the process of forming a shadow on the eyes of insects, and the working

principle of optical instruments and 4:11 Create reports the results of investigations on the formation of a shadow on mirrors, lenses and optical instruments.

TABLE 1. Component and indicator of evaluation instrument CIPP models

No	Aspect	Component	Indicator
1.	*Context*	The conformity of assessment instruments projects with KI and KD	Assessment instrument which present according to main competences in Curriculum 2013 The materials in assessment instrument has been coordinated with basic competence Indicators in assessment instruments are relevant with KD achievement which are used before.
2.	*Input*	Teacher's competence	Pedagogic Competence Understan students' characteristics Communicate effectively, emphatic and friendly to students Perform assesment and evaluation on learning process Personality Competence Act according to religion norm, religion, law, and Indonesian culture Show himself as honest and righteous personal Professional Competence Master the knowledge materials and concept which support the subject taught Master the standar competence and basic competence for the subject taught Develop the learning materials taught
		Learning Environment	There are facilities to support the assessment project There are Science labs There are tools for experiment which can be used to measure poluted environment level
3.	*Process*	Instrument carried out	On the beginning of lesson, teacher informs the studnets about the assessed aspect and criteria to achieve Teacher informs the students about the procedure and kind of assessment used Students learnt the lesson which baseed on LKPD with Optical on eyes theme Teacher assess students' skills using observation activities Assessment instrument project is filled based on the observation result
		Students' response toward lessons which apply LKPD Optical on Eyes theme	Students' response toward the lesson using LKPD on theme Optical Eyes Students' personality which is appeared when implementing lesson using LKPD Optical Eyes theme Develop problem solving skills Students work in group Dare to ask questions to teacher and friends
4.	*Product*	The result of implementation assesment instrument project	Assessment instrument project can map students' problem-solving skills Students' interest to follow / accept the lesson Materials understanding level Assessment instrument project contributes the decision on completion of student's learning based on KKM basic competence

Determine the Instrument's Specifications

CIPP models form instrument chosen in this study is the observation sheet. By using observation sheet CIPP instrument models can be done thoroughly, as observers directly observe the project assessment activities of learners in science learning. Determined indicators of four components, namely the context, input, process, and product. The grating of this instrument is based on the study of the components on the CIPP evaluation model that context, input, process, and product. Then the grating is translated to a corresponding point statement.

Writing Instruments

Writing CIPP model evaluation instrument based on the grating that has been made, next item is prepared on statement form. The grating includes components context, input, process, and product. The writing instrument of evaluation considers aspects of material, construction and the language used in order to use the evaluation instrument which is easy to understand and appropriate to provide an evaluation.

Determine the Scale of The Instruments and Scoring System

The scale of evaluation instruments development used is in the form of a scale of 1 to 4. The scoring system used in this study is the score acquisition which refers to the scale used is the scale of 1 to 4 based on the emergence of a large selection of available observations for each item given by the observer.

Reviewing Instruments

The study was conducted by the evaluation instrument experts to the developed area. CIPP model evaluation instrument development is validated by seven validators consist of two professors' experts, two Science teachers and three colleagues. Two expert lecturers consist of subject experts and experts in the field of evaluation. They are lecturers in Educational Evaluation Research study program. These experts will examine and give feedback on the coverage of the indicator into the grain of evaluation instruments that have been prepared. Experts in the field of measurement and instrument development education examine and advise on the substance, constructs and language in the instrument developed.

Result of Evaluation Instrument Validation

Feasibility CIPP model evaluation instrument was assessed in terms of substance, construction, and language. Score assessment from 7 validator then analyzed using a formula to calculate content Aiken's validity coefficient (V). Furthermore, V Aiken's figures confirmed the figures in Table V limit Aiken's for the number of category 4 votes, with 7 appraiser is 0.78 [24]. The calculation result of the value V of the 7^{th} validator on every aspect have good criteria. V value between 0.86 to 1 proves that each item has a valid validity. Based on the analysis of the value V of CIPP model evaluation instrument on the application of Science teaching project assessment indicates that the value of V is above the minimum value of V Aiken's. Consequently, the evaluation instruments meet the validity of the content. The validation result are shown in Table 2.

Criteria of evaluation instruments model CIPP has good reliability if the ICC ≥ 0.60, and the coefficient $\alpha \geq 0.70$. Instrument test results as presented in Table 3. The conclusion that the estimated coefficient of reliability of the instrument denfan both categories. That is, the CIPP model of evaluation instruments have been good consistency between inter-rater.

Revisions in this study conducted twice, by (1) the input and suggestions of the validator and (2) weaknesses and shortcomings while testing. Suggestions and feedback is used to improve the product so worthy CIPP model evaluation instrument used. The revision of the CIPP model of evaluation instruments : (a) improvement in writing some words are not in accordance with the rules of Indonesian true; (b) improvements in the language selection is less effective; (c) improvements to the point statement on the sheet CIPP model evaluation instrument on the order of items on several indicators; (d) improvements in grain that is not formulated statement brief, clear, and unequivocal; (e) improvements in point declaration contains statements that are not needed; (f) improvement of scoring rubric is less clear, and (g) the grain refinement are still experiencing looping statement.

TABLE 2. Result of contruct validity by using aiken's v formula

Instrument	Items	V	Criteria
Context	1,3,5,6	0,95	Valid
	2,4,7	1,00	Valid
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	Valid
	3a, 5b, 7d	0,86	Valid
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	Valid
	6a, 6d, 7a, 10d	0,86	Valid
Product	1, 3, 4, 6	1,00	Valid
	5, 9	0,90	Valid
	2, 7, 8	0,86	Valid

TABLE 3. ICC coefisien value interclass correlation coefisient

	Intraclass Correlationb	95% Confidence Interval		F Test with True Value 0			
		Lower Bound	Upper Bound	Value	df1	df2	Sig
Single Measures	,460a	,191	,789	6,102	8	40	,000
Average Measures	,836c	,836	,957	6,102	8	40	,000

CONCLUSION

Conclusion of this study are as follows: (1) the procedure of the evaluatioan instrumen CIPP model development follow the stages of research and development. (2) The quality of the developed products the development an evaluation instrument CIPP model has a valid criterion as an instrument, in terms of aspect of the construct, substance, and languange. All these aspects meet a very good criterion and can be used with revisions.

ACKNOWLEDGMENT

The authors thank to the Post Graduate Science, YSU, due to the funding of Direktorat Riset dan Pengabdian Pada Masyarakat (DPRM), Dirjen Dikti, Kemristek Dikti, SKIM Penelitian Hibah Pascasarjana for this work.

REFERENCES

1. Treagust, David F. 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. (2006).
2. Rosana, D. *Evaluasi Pembelajarn Sains Asesmen Pendekatan Saintifik Pembelajaran Terpadu* . (Yogyakarta. 2014).

3. Kuswanto, H. Penilaian Pembelajaran IPA. *Makalah disampaikan pada workshop Evaluasi dan Penilaian SMA N 3*, hal. 1. 2008.
4. Yahaya, A. H. The using of Model Context, Input, Process and Products (CIPP in :earning Programs Assessment). *Inrenational Conference on Challenges and Prospects in Teacher Education* (Johor: Universiti Teknologi Malaysia. 2001).
5. Sterling, D. Assessing Understanding. *Journal Science Scope*, **28** (4) p.33-37. (2005).
6. Callison, D. *AUTHENTIC ASSESSMENT.Scholl Library Media Activities Monthly 14 no 5.* dari http://www.ala.org/aasl/sites/ala.org.aasl/files/content/aaslpubsandjournals/slr/edchoice/SLMQ Authentic Assessment InfoPower.pdf. (1998).
7. Akinoglu, O. Assessment of the inquiry-based project implemetation process i science education student point of view. *Intrnational Journal Of Instruction*, 1-12. (2008).
8. Wiyarsi, A. *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%20authentic%20assessment.pdf
9. Sukmasari, V. P. *Pengembangan Instrumen Project Assesment Untuk Mnegukur Keterampilan Pemecahan Masalah.* (Yogyakarta. 2016).
10. Arikunto, S. *Dasar-dasar Evaluasi Pendidikan.* (Jakarta: PT. Bumi Aksara. 2011).
11. Nofiana, M. Pengembangan Instrumen Evaluasi Two-Tier Multiple Choice Question Untuk Mengukur Keterampilan Berpikir Tingkat Tinggi Pada Materi Kingdom Plantae. *Jurnal Inkuiri*, 60-74. 2014
12. Sanjaya, M. E. Pengembangan Instrumen Evaluasi pada praktikum Uji Enzim Katalase di SMA Negeri Titian Teras Muaro Jambi. *Edu-Sains* **4**, 2. (2015).
13. Arifin, Z. *Evaluasi Pembelajaran.* (Bandung: PT Remaja Rosdakarya . 2013).
14. Stufflebeam, D. *Systematic Evaluation "A Self-Intruction Guide to Theory and Practice".* (Hingham Massachusetts: Kluwer Academic Publishers. 1985).
15. TenBrink, T. D. *Evaluation A Practical Guide for Teacher.* (Columbia: McGraw-Hill. 1974).
16. Hasan, H. *Evaluasi Kurikulum.* (Bandung: PT Remaja Rosdakarya. 2008).
17. Safruddin, S. A. *Evaluasi Program Pendidikan.* (Jakarta: PT Bumi Aksara. 2014).
18. Guili, N. Z. 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,* **15** (4), 57 (2011).
19. K.Hakan & F. Seval, " CIPP Evaluation Model Scale: Development, Reliablity and validity,"*Procedia and Social and behavioral Science,* **15***, 592-599,*(2011).
20. G. Zhang, *et al.,*" Using the context, imput, process and product evaluation model (CIPP) as a comprehensive framework to guide the planning, implementation and assessment or service-learning program,"*Journal of Higher Education Outreach and Engagement,* **15**(4), pp. 57-84, (2011).
21. D.L., Stufflebeam, "The relevance of the CIPP Evaluation Model for Educational Accountability,"*Annual Meeting of the American Association of School Administrator,* 1971. Available at://eric.ed.gov/PDFS/ED062385.pdf
22. Oliva, P. *Developing The Curiculum Third Edition.* (New York: Harper Collins. 1992)
23. Nor, H.,Md., *A Reliability and Validity of an Instrument to Evaluate the School-Based Assessment System: A Pilot Study. International Journal of Evaluation and Research in Education (IJERE), Volume 5, No.2,148-157* (2016).
24. Nicholson, T., Using the CIPP Model to Evaluation Reading Instruction. *J.of Reading,* **32** (4) 312-318I (1989).
25. Stufflebeam, D.L., *The CIPP Model for Program Evaluation*. In:Madaus, G., Scriven, M. And Stufflebeam, D.(Eds), Evaluation Models: Viewpoints on Educational and Human Service Evaluation. Boston, MA: Kluwer-Nijhpff Publishing, 117-141 (1983).
26. Doyle, J.M. and Wood, W.C., Principles course assessment, accreditation, and the depreciation of economic knowledge. *J. of Educ. for Business*, **80**, 3, 165-171 (2005).
27. Crunkilton, C. *Curriculum development in Vocational and Technical Education.* Boston: Allyn and Bacon Inc. (1979).
28. Thiagarajan, Sivasailam, Semmel, Dorothy S., Semmel, Melvyn I. *Instructional Development for Training Teachers of Exceptional Children*. Bloomington, Indiana: Indiana University. 1975
29. Rusilowati, Ani. *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.(2013)
30. Aiken, L.R. *Three Coefficient for Analyzing the Reliability and Validity of Rating. Educational and Psychological Measurement*, **45**, 131-142 . (1985).