Developing evaluation instrument based on CIPP models on the implementation of portfolio assessment

Feni Kurnia, Dadan Rosana, and Supahar

Citation: AIP Conference Proceedings 1868, 080003 (2017);
View online: https://doi.org/10.1063/1.4995187
View Table of Contents: http://aip.scitation.org/toc/apc/1868/1
Published by the American Institute of Physics

Articles you may be interested in

Development of the evaluation instrument use CIPP on the implementation of project assessment topic optik
AIP Conference Proceedings 1868, 080006 (2017); 10.1063/1.4995190

Fostering a student's skill for analyzing test items through an authentic task
AIP Conference Proceedings 1868, 080002 (2017); 10.1063/1.4995186

The effectiveness of science domain-based science learning integrated with local potency
AIP Conference Proceedings 1868, 080001 (2017); 10.1063/1.4995185
Developing Evaluation Instrument Based on CIPP Models on The Implementation of Portfolio Assessment

Feni Kurnia\(^1\), a), Dadan Rosana\(^2\), b), Supahar\(^3\), c)

\(^1\)Graduate Student of Science Education, Graduate School of Yogyakarta State University, Indonesia
\(^2\)Department of Science Education, Graduate School of Yogyakarta State University, Indonesia
\(^3\)Department of Physics Education, Graduate School of Yogyakarta State University, Indonesia

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

Abstract. This study aimed to develop an evaluation instrument constructed by CIPP model on the implementation of portfolio assessment in science learning. This study used research and development (R & D) method; adapting 4-D by the development of non-test instrument, and the evaluation instrument constructed by CIPP model. CIPP is the abbreviation of Context, Input, Process, and Product. The techniques of data collection were interviews, questionnaires, and observations. Data collection instruments were: 1) the interview guidelines for the analysis of the problems and the needs, 2) questionnaire to see level of accomplishment of portfolio assessment instrument, and 3) observation sheets for teacher and student to dig up responses to the portfolio assessment instrument. The data obtained was quantitative data obtained from several validators. The validators consist of two lecturers as the evaluation experts, two practitioners (science teachers), and three colleagues. This paper shows the results of content validity obtained from the validators and the analysis result of the data obtained by using Aikens’ V formula. The results of this study shows that the evaluation instrument based on CIPP models is proper to evaluate the implementation of portfolio assessment instruments. Based on the experts’ judgments, practitioners, and colleagues, the Aikens’ V coefficient was between 0.86-1.00 which means that it is valid and can be used in the limited trial and operational field trial.

INTRODUCTION

Assessment system applied in curriculum 2013 have brought some significant changes to the process of assessing student’s learning outcomes. Previously, assessment had been done only at the end of the learning activity, but the implementation of curriculum 2013 changes the assessment system to be a measure all aspects of learning all along the learning activity. An effective assessment should be related to the way of learning, and the results can be used to inform the learning outcomes [1]. In addition, assessment is an integral part of the learning process [2]. The assessment process includes the collection of information through several assessment techniques and the decision making based on the learning outcomes. Assessment is a continuous process applied to determine the method proposed in lessons by teachers and supervisors, and at the same time it can give guidance for them to overcome the existing shortage [3]. Based on these opinions, it can be shown that the assessment is a continuous process that aims to make a decision based on the collection of information as well as to determine the achieved learning outcomes. This is in line with the curriculum assessment standards in curriculum 2013 that is required in Permendikbud (Education and Culture Ministerial Decree) No. 23 of 2016 [4], that reads "The assessment is the process of collecting and processing information to measure the achievement of students learning outcomes. Objective assessment of learning outcomes by educators aims to monitor and evaluate the process, the learning progress, and improvement of students learning outcomes on an ongoing basis. Assessment of students learning outcomes in primary education and secondary education includes aspects, attitudes, knowledge and skills. Skills assessment is an activity undertaken for measuring the learners’ ability to apply the knowledge in performing certain tasks. Skills assessment is done through practice, product, project, portfolio, and/or other techniques according to the competencies rated. The scoring system applied in curriculum 2013 is called authentic assessment. According to Permendikbud number 104 [5], authentic assessment can be defined as a form of assessment that requires students to show attitude, using the knowledge and skills gained from learning to apply in the actual
situation. Wiggins [6] stated that an assessment is called authentic when we directly examine student’s performance on appropriate intellectual tasks. In addition, process of authentic assessment can also be used to measure the 21st century skills of the students, one of which is the critical thinking skills. Referred from permenidikbud, the forms of authentic assessment are including assignment to the field, portfolios, projects, products, journals, laboratory work, and performance, as well as self-assessment [5]. One of authentic assessments applied in schools is a portfolio assessment.

A student portfolio is a collection of student’s work and related material that depicts student's activities, accomplishments, and achievements. The collection should include evidence of student’s reflection and self-evaluation, guidelines for selecting the portfolio contents, and criteria for scoring the quality of the work [7]. Sukmasari, et al [8] stated that the portfolio is a collection of artifacts or works of students during the learning process, which is then documented properly and regularly. Portfolios can take form as student's answers to the question asked by teacher, teachers’ observation notes, results of students’ interview, the concept maps and mind maps. Lubis [9] added that in the implementation of portfolio assessment, the assessment allows feedback, where the student’s portfolio can be analyzed collaboratively. Wulan’s research [10] showed that assessment of the portfolio has the ability to reveal the progress of students' learning process, learning attitude, interest and motivation, skills and misconceptions. In line with these results, the study by Rahmi, Y.L. and Alberida, H. [11] also stated that the assessment of the portfolio increases high-order thinking skills (HOTS).

The proper implementation of a portfolio assessment needs to be done by using the appropriate instrument of evaluation. The instrument of evaluation itself should be based on the assessment standards. Referring to this, Esmyiyati has conducted a research and development that resulted a valid and reliable portfolio assessment instruments [12]. In its development, the new instrument was tested on a limited scope, so it had not been able to be classified as a standardized instrument. Thus, the instrument needs to be applied to a broader scope so it can be used as a standardized instrument of portfolio assessment.

The proper implementation of an assessment instrument must necessarily be followed by an evaluation. The evaluation process is useful to establish the basic terminology of evaluation which is a systematic process that includes collecting, analyzing, and interpreting information to determine the extent to which learning objectives have been achieved by students [12]. The evaluation process can also be used as a guide to implement a plan/program [13,14]. There are four basic principles of a sound evaluation. That basic principles including utility, propriety, feasibility, and accuracy [15]. Evaluation activities should be done by using a good instrument of evaluation. A good instrument of evaluation is an instrument which able to provide an overall picture of the program to be evaluated and to produce evaluation outcomes according to the purpose of the evaluation. Instrument of evaluation used must meet the characteristics of a good instrument. To meet these criteria, the basic thing to do is a needs analysis. Through the analysis of the needs, the evaluator will gain clarity about the problems of the evaluated program.

A minimum requirement of a program consists of three dimensions, namely input, process and output. One of the evaluation models is CIPP. It has four components including that three-dimensional program. The terms CIPP stands for context, input, process, and product. Thus, the components of CIPP model evaluation represents the dimensions required by the program, especially in the implementation of portfolio assessment in science learning in Junior High School.

CIPP Evaluation Model first developed by Daniel Stufflebeam[16]. According to Stufflebeam, the CIPP evaluation model can provide an overview, results and provides useful information for consideration in making a responsible decision. In the CIPP evaluation model, there are four types of evaluations related to each other, namely:

1) Context Evaluation is useful in the early activities of the development program, which is to identify needs and design a program rationality. Questions that may be developed in the context of the evaluation in this study is whether there is compatibility between the portfolio assessment instruments with core competence and basic competences pressure material?

2) Input Evaluation provides information on the resources required for the proposed program. In this case, it is the implementation of portfolio assessment instruments. Questions that can be developed are:
   a) How well does the teacher understand the portfolio assessment?
   b) Are teachers skilled in the implementation of the portfolio assessment in the assessment process?
   c) How is the learning environment of students in the school?

3) Process Evaluation is very closely associated with learning and it is focused on how effectively the effects of the implementation of the system or program evaluated [17]. Questions that can be developed in this evaluation are:

080003-2
a) How is the implementation of the portfolio assessment instrument in the learning process in the school?
b) How is the response of students to the portfolio assessment instruments?

4) Product Evaluation is the final evaluation of the program. It decides whether the program can be continued, modified or terminated. The program in this research is the implementation of portfolio assessment instruments. Questions that can be developed in the evaluation of the results are:

a) How are the results obtained from the use of portfolio assessment instruments?
b) What follow-up should be done about the results of the implementation of portfolio assessment instruments?

CIPP model is used as a basis for the development of evaluation instruments. It is considered suitable to be used to obtain data on evaluation of the implementation of the portfolio assessment. CIPP evaluation model has a meaning in the matter of the assessment planning, in the process of the assessment implementation, as well as the final outcomes obtained from the implementation of the portfolio assessment.

In performing the evaluation of the implementation of portfolio assessment, it will be done based on the CIPP evaluation model. This evaluation model is suitable to use considering that CIPP evaluation model sees a program as a system. Thus, the evaluation of the program as the system should be done in details according to the components [18].

The components of the system will be evaluated with CIPP models are:

1. Context
   Component context is the stage where an evaluator identify several factor, such as teacher, working facilities, student, work atmosphere, the role of the school committee, the community, as well as several other factors that have an effect on the possibility of the system being evaluated [19]. Furthermore, according to the context, component is closely related to the target population, the analysis of the needs of the institution, and the opportunity to realize the needs of the institution [20]. In addition, an evaluator assigns the situation where the system will be evaluated and perform an analysis of unmet needs, as well as identification of any reason why the background of those requirements have not been or can not be reached. Evaluation in this context section focuses on the evaluation of activities related to the analysis of needs, needs that have been achieved or not achieved. The analysis also examines more details by looking for reasons of those needs fulfillment. In addition, to determine the needs of the program, this context evaluation may also specifies the program objectives.

2. Input
   Input evaluation aims to organize a decision to be taken, to specify alternatives about to be taken, and how to achieve the unmet needs through appropriate work procedures [20]. The same is expressed by Stufflebeam [21] who stated that the evaluation in this section helps to determine the information that will be used to meet the objectives or needs. In this regard, the input evaluation is closely related to the determination of the source and any strategy that will be used to achieve the purpose of the system or program being evaluated [17]. Examples of factors that affect the struggle in achieving the goal are the teachers’ way of teaching, the use of instructional media and the learning environment.

   Based on the description from some experts 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 can not be reached. The strategy may come from teachers’ way of teaching through teaching skills and the use of media in learning process, and from students in the form of spirit, concentration, and students understanding.

3. Process
   Process evaluation is related to the reciprocal arisen from the implementation of a system or program which is being evaluated [21]. Furthermore, it is also known that the evaluation process is very closely associated with learning. Evaluation process is focused on how effectively the effects of the implementation of the system or program being evaluated [17].

   The description above explains that the evaluation process is based on several experts. So, it can be seen that the evaluation process is evaluation activities that focus on the implementation of a system or program which is being evaluated.

4. Product
   Product evaluation or the result has the aim to know the results of what has been achieved from the implementation of the system or program, and to follow up what will be done after the system or specific programs implemented. Information obtained from the evaluation of these products is very important for the evaluation results obtained to determine the further follow-up to be taken [17].

   Products evaluation is related to the analysis of the results of the implemented instrument. The information obtained can show the results of what has been achieved. So, the information can be determined with the next steps
to be taken to follow up the implementation of the systems or programs in the future. Nevertheless, this paper will mainly focus in content validity of the instrument of evaluation that had been developed.

METHOD

Type of Research

This study used research and development (R & D) method; adapting 4-D by the development of non-test instrument. Evaluation instrument that will be developed is an instrument of CIPP model-based evaluation. CIPP is an evaluation model which consists of Context, Input, Process, and Product. Context presented the entire data preparation of portfolio assessment instruments implementation. Input of the implementation of portfolio assessment instruments were obtained from a given input and impact from implementing instruments. Process was carried out by observing the implementation of portfolio assessment instruments. While the product in this study saw the objectives achievement of portfolio assessment instruments. The procedure of CIPP model evaluation instrument development on the implementation of portfolio assessment includes steps as follows: 1) conducting the preliminary study, 2) determining the specifications instrument, 3) developing an evaluation instrument, 4) determining the scale of instruments, 5) determining the system of scoring, 6) reviewing of evaluation instruments, 7) assembling an evaluation instrument, 8) trying out, 9) analysing the test results, 10) fixing the instrument, 11) carrying out measurements, 12) interpreting measurement results [23].

Data Analysis Techniques

The analysis of content validity is done by using Aiken’s V analysis. There are several steps of the analysis on the data obtained from the validators, they are 1) Tabulating all the data obtained from the validation result. 2) Calculating the content validity coefficient by using Aikens’ equation. The Aikens’ V coefficient ranges from 0-1. 3) Comparing the results of the Aikens’ V coefficient calculation with the V’aiken's category table. Items have good content validity and support the overall content validity when the minimum number is 0.86. The number 0.86 is the minimum boundary coefficient of Aikens' V using 4 rating categories and 7 raters [24]. The developed assessment instrument is considered feasible if it meets the validity of good content.

Analysis of content validity is based on descriptive and quantitative techniques. Quantitative analysis of this data uses Aiken’s V [22] by the following formula:

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

Description :

\(s = r - l\)  
\(n = \) number of panels of assessors  
\(l = \) lowest validity assessment  
\(c = \) highest validity assessment  
\(r = \) the numbers given by an assessor

RESULTS AND DISCUSSION

Developmental Procedure of CIPP Based Instrument of Evaluation

Development Procedure in this research is the integration of 4D model with the non-test instrument development model. The steps in the non-test instrument development model are sorted and combined in the 4D development model. The procedures of CIPP model evaluation instrument development in the portfolio assessment implementation in science learning to measure critical thinking skills are explained in the following paragraphs.

1. Conducting Preliminary Studies

Teacher competency analysis is conducted to obtain information on the understanding of teachers and how to use the portfolio by teacher assessment instruments. This analysis is conducted through interviews with some of the
Junior High School science teachers. Analysis of material is done by studying $KI$ (Core Competence) and $KD$ (Basic Competence) in Curriculum 2013. The competencies used in this study are $KD$ 3.8 and 4.8 that is about the liquid pressure.

2. Determining the Instrument’s Specifications

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

3. Writing instruments

Writing CIPP model evaluation instrument is based on the grating that has been made. the next item is prepared on a 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.

4. Determining the Scale of The Instruments and Scoring System

The scale of evaluation instruments development used is in a scale of 1 to 4. The scoring system used in this study is the score acquisition. It refers to the scale used that 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.

5. Reviewing Instruments

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

**Results for Evaluation Instrument Validation**

Content validity is one of three empirical validation procedures that needs to be done in developing an instrument [25]. In the process of the instruments validation, the instruments need to be consulted to the experts. In this case, there are seven experts including lecturers and practitioners. Validators give a check mark in the column available in the validation sheet, they also provide suggestions or input to the instrument being reviewed. Subsequently the expert and practitioner checklist results are converted to four scales for analysis using the Vaiken formula.

Feasibility of CIPP model evaluation instrument is assessed in terms of substance, construction, and language. Score assessment is obtained from 7 validators then analyzed using a formula to calculate Aiken's validity coefficient (V). Furthermore, V Aiken's figures are confirmed by the limit figures in Table V Aiken's for 4 rating categories with 7 raters is 0.86 [24]. Judgment of validation results by the experts are divided into four categories (Lynn, 1986: 384) namely: (1) the items are accepted, (2) the items are accepted but need to be revised, (3) the items are repaired, and (4) the items are discarded. The calculation result of the value V of the seven validators on every aspect shows good criteria. The results of V values are between 0.86 to 1. It proves that each item has met valid criteria. Based on the analysis of the value V of CIPP model evaluation instrument on the implementation of portfolio assessment in science learning, the results indicate that the value of V is above the minimum value of V Aiken's. Consequently, the CIPP model of evaluation instruments meet the validity of the content. The validation results are shown in table 1 below.
**TABLE 1. Calculation Results of Construct Validity by Using Aiken’s V Formula**

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Items</th>
<th>V</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Context</td>
<td>2,3,4,6</td>
<td>1.00</td>
<td>Valid</td>
</tr>
<tr>
<td></td>
<td>1,5</td>
<td>0.90</td>
<td>Valid</td>
</tr>
<tr>
<td>Input</td>
<td>1b, 1c, 1d, 2b, 2c, 2d, 3b, 3c, 3d, 4b, 4c, 4d, 5b, 5c,5d, 6b, 6c, 6d, 7c, 8a, 8b, 8c, 8d, 9a, 9b, 9c, 9d, 10b, 10c, 10d</td>
<td>1.00</td>
<td>Valid</td>
</tr>
<tr>
<td></td>
<td>5a, 6a, 7a,7b, 10a</td>
<td>0.90</td>
<td>Valid</td>
</tr>
<tr>
<td></td>
<td>1a, 2a,3a,4a,7d</td>
<td>0.86</td>
<td>Valid</td>
</tr>
<tr>
<td>Process</td>
<td>1d, 2a, 2b, 2c, 3a, 3b, 3c, 4a, 4b, 4c, 5a, 5b, 5c, 6a, 6b, 6c, 7a, 7b, 7c, 8a, 8b, 8c, 8d,9a, 9b, 9c, 10a, 10b, 10c, 10d</td>
<td>1.00</td>
<td>Valid</td>
</tr>
<tr>
<td></td>
<td>1a, 2d, 3d, 4d</td>
<td>0.90</td>
<td>Valid</td>
</tr>
<tr>
<td></td>
<td>1b, 1c, 5d, 6d, 7d, 9d</td>
<td>0.86</td>
<td>Valid</td>
</tr>
<tr>
<td>Product</td>
<td>2,4,6,7,8</td>
<td>1.00</td>
<td>Valid</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.90</td>
<td>Valid</td>
</tr>
<tr>
<td></td>
<td>1,5,9</td>
<td>0.86</td>
<td>Valid</td>
</tr>
</tbody>
</table>

**Revision of the Product**

Revisions are made to the content and appearance of products that aim to enhance CIPP model evaluation instrument on the implementation of the portfolio assessment measuring critical thinking skills on the liquid pressure theme. Revisions in this study are conducted twice, by (1) the input and suggestions from the validators and (2) weaknesses and shortcomings while testing. Suggestions and feedback are used to improve the product, so a proper CIPP model evaluation instrument can be used. The revisions of the CIPP model evaluation instruments are: (a) improvement in writing some words that are not appropriate with the rules of Indonesian language; (b) improvements in the language selection that is less effective; (c) improvements of the point statement on the CIPP model evaluation instrument sheet, on the order of items of several indicators and (d) improvements in item that is not stated in a brief, clear, and assertive way. Overall, this instrument of evaluation based on CIPP is valid and it is good be used as an evaluation instrument in the implementation of portfolio assessment in Junior High School.

**CONCLUSION**

Conclusion of this study are: (1) the procedure of the evaluation instrument of CIPP model development follows the stages of research and development. (2) The quality of the products development of the CIPP model evaluation instrument is in the valid criteria as an instrument, in terms of aspect of the construct, substance, and language. All of these aspects meet a very good criterion and can be used in the limited trial and operational field trial.

**ACKNOWLEDGMENT**

The authors thank to the Direktorat Riset dan Pengabdian Masyarakat, Direktorat Jendral Pendidikan Tinggi, and Kemristek Dikti for the funding of Pendanaan Penelitian dan Pengabdian Masyarakat di Perguruan Tinggi through the scheme of Penelitian Hibah Pascasarjana year 2017 for this work.

**REFERENCES**

3. Yahaya, A. H. The using of Model Context, Input, Process and Products (CIPP in Learning Programs