



## **Educational Program Evaluation using CIPP Model**

Warju

Department of Mechanical Engineering, Faculty of Engineering, Universitas Negeri Surabaya, Indonesia

### **ARTICLE INFO**

*Article history:*

Received 02 December 2015

Received in revised form 12

January 2016

Accepted 28 January 2016

Available online 01 February 2016

*Keywords:*

Evaluation

Education program

CIPP model

*Corresponding author:*

warjuunesa@gmail.com

### **ABSTRACT**

There are many models of evaluation that can be used to evaluate a program. However, the most commonly used is the context, input, process, output (CIPP) evaluation models. CIPP evaluation model developed by Stufflebeam and Shinkfield in 1985. The evaluation context is used to give a rational reason a selected program or curriculum to be implemented. A wide scale, context can be evaluated on: the program's objectives, policies that support the vision and mission of the institution, the relevant environment, identification of needs, opportunities and problems specific diagnosis. Evaluation input to provide information about the resources that can be used to achieve program objectives. Evaluation inputs used to: find a problem solving strategy, planning, and design programs. Evaluation process serves to provide feedback to individuals to account for the activities of the program or curriculum. The evaluation process is conducted by: monitoring sources can potentially cause failure, prepare a preliminary information for planning decisions, and explain the process that actually happened. Product evaluation measure and interpret the achievement of goals. Evaluation of the products also come to: the measurement of the impact of the expected and unexpected. The evaluation is conducted: during and after the program. Stufflebeam and Shinkfield suggest product evaluation conducted for the four aspects of evaluation: impact, effectiveness, sustainability, and transportability. The decision making process is done by comparing the findings / facts contained in context, input, process and product standards or criteria that have been set previously.

## 1. Introduction

### 1.1 Measurement, Assessment, and Evaluation

Speaking of the evaluation, we will be faced with three terms that are interrelated with each other even in the field sometimes occur overlapping in use, namely measurement, assessment, and evaluation. Actually, what the common thread of these three terms? Are these three terms together or is it different? If equal, the equation which is so well if different, the difference where? To answer these questions will be discussed with a review of the literature. Measurements used to collect information on which to base a decision, and evaluation using the measurement results to make a decision (Mehrens & Lehman, 1984). Meanwhile, according to Griffin and Peter (1991), measurement is the determination of the object observed numbers/data retrieval. The assessment is the interpretation of the results of observation and description of the results of the overall measurement. While evaluation is an activity use the assessment results as consideration for decision making. So, from both the above opinion can be concluded that measurement, assessment and evaluation have understanding and different functions, but between the one and the other interrelated.

In more detail, in the *Encyclopedia of Evaluation* stated that “*evaluation is an applied inquiry process for collecting and synthesizing evidence that culminates in conclusions about the state of affairs, value, merit, worth, significance, or quality of a program, product, person, policy, proposal, or plan*”(Mathison, 2005:140). From the above it can be seen that the evaluation is a process of inquiry that is used to gather and synthesize the evidence ended conclusions about the state of affairs, value, merit, feasibility, significance, or quality of the programs, products, people, policies, proposals or plans. The results of the evaluation involves two aspects, namely empirical and normative aspects. Therefore, the evaluation activities in contrast to the basic science research, clinical epidemiology, investigative journalism, or public poll.

Evaluation has a very broad sense and is not a new concept. Hadley and Mitchell (1994:48), defines evaluation as “*applied research carried out to make or support decisions regarding one or more service programs*”. While understanding the evaluation according to the United Nations Development Program/UNDP (2006:6)<sup>5)</sup> is “*selective exercise that attempts to systematically and objectively assess progress towards and the achievement of an outcome. Evaluation is not a one-time event, but an exercise involving assessments of differing scope and depth carried out at several points in time in response to evolving needs for evaluative knowledge and learning during the effort to achieve an outcome*”. Further evaluation is the process of delineating, obtaining and providing useful information for judging decision alternative. From the definition of the evaluation above it can be concluded that the evaluation is an activity to gather information about the workings of something, which then the information is used to determine an appropriate alternative in the decision.

### 1.2 Educational Evaluation

So what is the evaluation of education? Understanding evaluation of education, according to Law Number 20 Year 2003 on National Education System Article 1 stated that educational evaluation is control activities, underwriting, and determination of the quality of education to the various components of education at every track, level and type of education as a form of education provision<sup>7)</sup>. While accreditation is the program's eligibility assessment activities in education units based based on defined criteria. Criteria that refers to the eight National Education Standards, created by the National Education Standards Agency where the National Education Standards are the minimum criteria regarding the educational system in the whole territory of the Republic of Indonesia. Eight National Education Standards are: 1) the content standard, 2) the process standard, 3) the competency standard, 4) the teacher and staff standard and, 5) the facilities and infrastructure standard, 6) the management standard, 7) the financing standard, and 8) the assessment standard.

### 1.3 Monitoring and Evaluation

In the field of education is also often used the term monitoring and evaluation is commonly abbreviated as *monev*. Actually, what the common thread of the two terms? Do words have a striking difference? To answer these questions, we can find the answer from literature. According to the United Nations Development Program/UNDP (2006:6), monitoring is defined as “*a continuing function that aims primarily to provide the management and main stakeholders of an ongoing intervention with early indications of*

*progress, of lack thereof, in the achievement of results. An ongoing intervention might be a project, programme or other kind of support to an outcome*). While understanding the evaluation according to the United Nations Development Program/UNDP (2006:6) is *“selective exercise that attempts to systematically and objectively assess progress towards and the achievement of an outcome. Evaluation is not a one-time event, but an exercise involving assessments of differing scope and depth carried out at several points in time in response to evolving needs for evaluative knowledge and learning during the effort to achieve an outcome”*.

#### 1.4 Evaluation and Research

In the field of education is also often overlapping use of terms between evaluation and research. What is the common thread that can be drawn from both? Mertens (2009) states that *“while there is much overlap between the world of research and evaluation, evaluation occupies some unique territory”*. More detail, Trochim (2006) argues that *“evaluation is unique because of the organizational and political context in which it is conducted, which require skills in management, group processes, and political maneuvering that are not always needed in research”*. While Arikunto (2013) states that in terms of objectives, evaluation research, and evaluation of education, there are three differences. *First*, evaluation research is usually conducted in the context of decision making. *Secondly*, the evaluation is usually carried out with a limited purpose. *Thirdly*, evaluation studies with educational research related to consideration of the meaning and value.

In the research, researcher wanted to find a picture of something later described, whereas in program evaluation, the evaluator want to know how high the quality or condition of something as a result of the implementation of the program, after the data was collected as compared to specific criteria or standard. In the research, researcher led by the formulation of the problem, whereas in program evaluation, the evaluator would like to know the level of achievement of the program, and if the objectives have not been achieved, the evaluator wanted to know where the deficiencies and why. The results are used to determine the follow-up or a decision to be taken. There are many models of evaluation that can be used to evaluate a program. However, the most commonly used are Context, Input, Process, Product (CIPP) model.

## 2. CIPP Evaluation Model

The basic framework of the CIPP was complete (context evaluation to help develop goals, input evaluation to help shape proposal, process evaluation to guide implementation, and product evaluation to serve recycling decisions. This model has a basic framework that is complete, ie the evaluation context to help formulate objectives, evaluation input help in preparing the program, the evaluation process is to direct enforceability of a program, and product evaluation to determine the achievement of a program.



**Figure 1.** The CIPP model of evaluation  
Source: <https://www.google.co.id/search>

Dharma, et al (2013:345) states that the curriculum evaluation carried out for two interests, namely the measure of accountability and feedback curriculum implementation. Evaluation of curriculum accountability measures if the evaluation of the curriculum aims to examine accountability in achieving educational goals with regard to quality and quantity of yield and cost-effectiveness. The evaluation is conducted to provide feedback on the implementation of the curriculum when the evaluation aims to eliminate the limitations and continuously improve performance at both the individual and institution. CIPP is an acronym of context, input, process, and product. CIPP evaluation model has been widely used to evaluate a program, policy, and can also be used to evaluate the curriculum on a small scale. Things that need to be evaluated in each component of the activities are:

## 2.1 Context Evaluation

Context evaluation is used to give a rational reason a program or curriculum have to be implemented. On a large scale, can be evaluated in the context of: the program's objectives, policies that support the vision and mission of the institution, the relevant environment, identification of needs, opportunities and specific problems diagnosis. Need assessment is a common example of context evaluation. In a small scale, can be applied to evaluate the learning program objectives. Context evaluation to serve planning decision.

## 2.2 Input Evaluation

Input evaluation to provide information about the sources that can be used to achieve program objectives. Input evaluation is used to: find a problem solving strategy, planning and design programs. The results of the input evaluation: budget, schedule, proposals and procedures. In learning activities, input evaluation can also be done to find sources that can be used in the learning process so that it can serve to establish appropriate learning strategies. Input may contain: student component, infrastructure, media, teacher, etc. Input evaluation to serve structuring decision.

## 2.3 Process Evaluation

Process evaluation serves to provide feedback to the individual to take responsibility for the activities of the program or curriculum. Process evaluation is carried out with: monitor potential sources can cause failure, prepare a preliminary information for planning decision, and explain the process that actually happened. Process evaluation requires: data collection instruments (observation sheet, assessment scale, field notes, etc.). Process evaluation to serve implementing decision.

## 2.4 Product Evaluation

Product evaluation measure and interpret the achievement of objectives. Product evaluation also arrive at: measuring the impact of the expected and unexpected. The evaluation is conducted: during and after the program. Stufflebeam suggest the product evaluation conducted for four aspects of evaluation: *impact, effectiveness, sustainability, and transportability*. Product evaluation requires: an instrument (such as test sheet, interview sheet, and observation sheet) to observe behavior change after the implementation of the learning program. Product evaluation to serve recycling decision.

The three stages in the technical educational activities interrelated, namely: (1) input is community needs and resources, (2) process is about learning, and (3) output is workforce skills (Kumar 2011). With the CIPP evaluation model will be produced four types of decisions, namely: 1) decision in planning that affects the choice of goals and objectives, 2) make decision that ensure the design of strategies and procedures are optimal to achieve the goals, 3) implement the decision: the work done to bring and improve by selecting the designs, methods, and strategies, and 4) repeat the decision to establish continuity, change or terminate the program activities.

### 3. The Objective and Benefits of Program Evaluation

What exactly is the purpose of implementation of the program evaluation? Wholey, Hatry & Newcomer (1994:1) state that “*program evaluation attempts to provide processes that agencies of all kinds can apply to obtain better, more valid, answers to these questions*”. While Dharma, et al (2013) explains that all program activities using the program evaluation to see the level of success that has been achieved, determine the effectiveness and efficiency of ongoing program and to obtain information to establish the next program is worth doing.

From the above it can be concluded that the program evaluation is a unit or entity activities aimed at collecting information about the realization or implementation of a policy, takes place in a continuous process, and occurs in an organization involving a group of people to decision-making.

There are four benefits that can be drawn from the program evaluation activities, namely:

1. Termination of program.
2. Revise program.
3. Continuing program.
4. Disseminate program.

At least five traits into program evaluation requirements, namely:

1. Referring to the rules which apply.
2. Do systematically.
3. Identified determinants of success and failure of the program.
4. Using standard benchmarks.
5. The evaluation results can be used for decision making.

To produce the right decision from the program evaluation depends on the ability of the person doing the program evaluation (evaluator). There are at least six conditions to be evaluator, namely:

1. Able to carry out.
2. Accurate.
3. Objective.
4. Be patient and persevering.
5. Carefully.
6. Responsible.

### 4. Evaluation Execution

According Arikunto (2013:228), there are several steps that must be done by the evaluator at the time of the program evaluation, namely:

1. Evaluator conduct a bookreview, field, and gather information from experts to gain an overview of the issues to be evaluated.
2. Evaluator formulates the program/research problem in the form of evaluation research questions.
3. Evaluator develop a evaluation research proposal.
4. Evaluator set evaluation planning, arranging instruments, preparing human resources, and carry out trials instruments.
5. Implementation of the evaluation in a form the evaluation model that has been adapted.
6. Evaluator collecting data with instruments that have been prepared based on the details of the components to be evaluated.
7. Analyze the collected data by comparing it to benchmarks/criteria has been established in accordance with the objectives set by the program manager.
8. Summing up the results of the evaluation based on an idea of the extent of the data in accordance with the benchmarks/criteria.
9. Information on the results of the evaluation study submitted to the program manager or the parties ask for help to the evaluation researcher.

According to Wholey, Hatry and Newcomer (1994:233-385), there are some procedures/methods for data collection that can be done in the implementation of program evaluation, namely:

1. Use of rating by trained observers.
2. Designing and conducting surveys.
3. The use of expert judgment.
4. The use of role-playing in evaluation.
5. The use of Focus Groups Discussion (FGD).
6. Managing field data collection from start to finish.
7. Collecting data from agency records.

Data analysis in the program evaluation can use some methods/approaches, namely:

1. Using statistics appropriately.
2. Using regression models to estimate program effects.
3. Cost-benefit analysis.

Further according to Wholey, Hatry and Newcomer (1994: 457)<sup>17)</sup>, steps in conducting a cost-benefit analysis, there are three steps, namely: (1) determine the benefits of a proposed or existing program and place a dollar value on those benefits, (2) calculate the total cost of the program, and (3) compare the benefits and the costs.

## 5. The Decision of Program Evaluation

In general, there are three follow-up/recommendations made by the evaluator to the decision maker to a running program, namely:

1. The program continued and disseminated because they were considered good and successful.
2. The program revised because there are things that are not in accordance with the benchmarks/criteria desired.
3. The program stopped because there is less evidence/not good or there was a violation in the implementation.

## 6. Conclusion

There are many models of evaluation that can be used to evaluate a program. However, the most commonly used are context, input, process, product (CIPP) evaluation model. CIPP evaluation model developed by Stufflebeam and Shinkfield in 1985. The purpose of the evaluation is to collect data or information to be compared with the criteria that have been made and then conclude. The gap between the real condition in the hope that is sought. The gap obtained a description of whether the implementation of the programs surveyed was appropriate, less appropriate, or not in accordance with predetermined criteria. The decision making process is done by comparing the findings/facts contained in context, input, process and product with a standard or predefined criteria.

## Acknowledgment

The author would like to thank the Vice Rector for Academic Affairs of the State University of Surabaya Dr. Hj. Yuni Sri Rahayu, M.Si which has funded the publication of this scientific article/paper.

## References

- Arikunto, Suharsimi. (2013). Manajemen penelitian. Jakarta: PT. Rineka Cipta.
- Ajit Kumar, N. S. (2011). Criteria for evaluating the quality of a question paper [printed version]. Journal of Technical Education and Training, 3, 59-64.

Dharma, et.al. (2013). Tantangan guru SMK abad 21. Jakarta: Direktorat Pembinaan Pendidik dan Tenaga Kependidikan Pendidikan Menengah, Direktorat Pendidikan Menengah, Kementerian Pendidikan dan Kebudayaan.

Government Regulation Number 19 Year 2005 on National Education Standards.

Griffin, P., & Peter, N. (1991). Education assessment and reporting. Sidney: Harcourt Brace Javanovich Publisher.

Hadley, R. G., & Mitchell, L. K. (1994). Counseling research and program evaluation (1<sup>st</sup> Ed.). California: Brooks Cole.

Law Number 20 Year 2003 on National Education System.

Mertens, D. M. (2009). Transformative research and evaluation. New York: Guilford.

Mehrens, W. A., & Lehman, J. J. (1984). Measurement and evaluation: in education and psychology. New York: Holt, Rinehart and Winston Inc.

Mathison, S. (Ed.). (2005). Encyclopedia of evaluation. Thousand Oaks, California: SAGE Publications, Inc.

Stufflebeam, D.L. & Shinkfield, A.J. (1985). Systematic evaluation: a self-instructional guide to theory and practice. Kluwer-Nijhoff Publishing.

Trochim, W. M. (2006). The research methods knowledge base (2<sup>nd</sup> Ed). Ithaya, New York: Author. (<http://www.socialresearchmethods.net/kb/>), rethieved April 1<sup>st</sup>, 2015.

The CIPP model of evaluation.

(2015). ([https://www.google.co.id/search?hl=en&site=img&tbm=isch&source=hp&biw=1056&bih=505&q=cipp+model&oq=CIPP+M&gs\\_l=img.3.0.0i5j0i8i30j0i30i3j0i24.2800.4381.0.8166.6.6.0.0.0.0.396.1009.2-1j2.3.0....0...1ac.1.64.img..3.3.1008.OBdj30okL8Q](https://www.google.co.id/search?hl=en&site=img&tbm=isch&source=hp&biw=1056&bih=505&q=cipp+model&oq=CIPP+M&gs_l=img.3.0.0i5j0i8i30j0i30i3j0i24.2800.4381.0.8166.6.6.0.0.0.0.396.1009.2-1j2.3.0....0...1ac.1.64.img..3.3.1008.OBdj30okL8Q)), rethieved July 7<sup>th</sup>, 2015.

The CIPP evaluation model. (2015).

([http://www.cglrc.cgiar.org/icraf/toolkit/The\\_CIPP\\_evaluation\\_model.htm](http://www.cglrc.cgiar.org/icraf/toolkit/The_CIPP_evaluation_model.htm)), rethieved October 16<sup>th</sup>, 2015.

United Nations Development Programme (UNDP). (2006). Handbook on monitoring and evaluating for results. New York: Author.

Worthen, B. R. & Sanders, J. R. (1973). Educational evaluation: theory and practice. Belmont, California: Wadsworth Publishing Inc.

Wholey, J. S., Hatry, H. P., Newcomer, K. E. (1994). Handbook of practical program evaluation. San Francisco: Jossey-Bass Publishers.