
DISCUSSION

EVALUATION MODELS IN EDUCATIONAL PROGRAM: STRENGTHS AND WEAKNESSES

Vo Thi Kim Anh*

PhD. Student, Universiti Malaysia Sabah

English Department, University of Foreign Language Studies, The University of Danang

131 Luong Nhu Hoc, Khue Trung, Cam Le, Danang, Vietnam

Received 5 November 2017

Revised 21 March 2018; Accepted 29 March 2018

Abstract: In the 21st century, evaluation in education has been paid great attention and the evaluation models in education which were created in the 20th century have been further developed and widely applied in educational evaluation. The paper provides readers with comprehensive discussions on the four well-known evaluation models in education: Tyler's objective model, Stake's responsive model, Scriven's goal free model and Stufflebeam's CIPP model. These models have a long history and have been thoroughly developed over time. The application of these four models is found in many fields of evaluation, but mostly in educational program evaluation. In order to help educational evaluators have better and deeper understandings of the four models, the paper presents the nature of the models, the characteristics of the models, as well as discusses strengths and weaknesses of each model.

Keywords: educational evaluation, evaluation model, objective

1. Introduction

Educational evaluation includes a wide array of activities like student assessment, measurement, testing, program evaluation, school personnel evaluation, school accreditation, and curriculum evaluation. The term "evaluation" is sometimes used ambiguously in relation to other terms 'assessment and testing'. However, evaluation does refer to the same thing as "assessment and testing" even though assessment instruments such as tests can be made use of in evaluation. Evaluation is regarded as "the systematic attempt to gather information in order to make judgments or decisions" (Lynch, 1996, p.2).

Evaluation is more thoroughly defined as "the process of delineating, obtaining, providing, and applying descriptive and

judgmental information about the merit and worth of some object's goals, design, implementation, and outcomes to guide improvement decisions, provide accountability reports, inform institutionalization/dissemination decisions, and improvement decisions, and understanding of the involved phenomena" (Stufflebeam, 2003, p.34).

Educational evaluation was first developed in the USA, and then spreads widely to all parts of the world (Madaus & Stufflebeam, 2000). Madaus and Stufflebeam (2000) divides the history of evaluation in education, which dates back to 150 years ago into seven different periods: *Age of Reform (prior to 1900)*, *Age of Efficiency and Testing (from 1900 to 1930)*, *Tylerian age (from 1930 to about 1945)*, *Age of Innocent (from 1946 to 1957)*, *Age of Development (from 1958 to 1972)*, *Age of Professionalization (from 1973*

* Tel.: 84-905190029

Email: vokimanh74@yahoo.com

to 1983), Age of Expansion and Integration (1983 to 2000).

A large number of evaluation models were created in the 40s, 50s and 60s. Among them are the four models: Tyler's objective model, Stake's responsive model, Scriven's Goal Free model and Stufflebeam's CIPP model, which have been developed and widely applied in educational evaluation in general and educational program evaluation in particular.

2. Tyler's objective model

Ralph W. Tyler plays an essential role in the development of educational evaluation and testing and his name is given to the period of evaluation development from 1930 to 1945 (Madaus & Stufflebeam, 2000). He is the father of the objective model, which is called Tylerian model or Tyler's objective model, which was first created in the 1940s. This model has the basic principles involving matching the pre-behavioral objectives with the actual outcome (Tyler, 1949). Evaluation is conceptualized in the view of Tyler as a comparison of intended outcomes with actual outcomes. In fact, as suggested by the name, Tyler's model is based on the objective-oriented theory. The model considers curriculum as a means of aiming toward an educational object.

2.1. The nature and characteristics of Tyler's objective model

The nature of Tyler's objective model is that it evaluates the degree to which an instructional program's goals or objectives were achieved. The model mainly involves the "careful formulation according to three educational goals (the student, the society, and the subject matter) and two goal screen (a psychology of learning and a philosophy of education)" (Popham, 1995, p.25). The result goals are then transformed into measurable objectives.

With Tyler's evaluation, the evaluator can determine the level to which the objectives

of the program are achieved. Unattained objectives mean that the instructional program has inadequacies. By contrast, attained objectives show successful instructional education program. However, as the objectives can be changed during the implementation of the program or the program may not have clear objectives, Tyler's objectives model can be only used to evaluate those with clear and stable objectives.

Even though Tyler's objective model was first created for use in evaluating educational programs, objective-based evaluation can be found in all kinds of services and it is usual to see the government requirements specifying that evaluations should be carried out to determine the extent to which each funded program achieved its objectives (Stufflebeam, Coryn, & Chris, 2014). The objective-based approach developed from Tyler's objective model is especially applicable in evaluating tightly focused programs that have clear and supportable objectives. Such evaluations can be strengthened by assessing program objectives against recipients' assessed needs, looking for side effects, and studying the process together with outcomes (Scriven, 1974, 1991).

If the evaluator wishes to identify the level the program meets its intended objectives, Tyler's objectives model is the best choice. Yet, the model can only be used if the program has clear objectives. Therefore, before deciding to use Tyler's objective model, the evaluator must make sure that the objectives of the program evaluated are clear and stable during the implementation of the program.

Tyler (1976) focused the curriculum development on four areas: selecting educational purposes, selecting learning experiences, organizing learning experiences and evaluation. The principles of Tyler's model engage the answers to the questions: "What should the educational objectives of

the curriculum be? What learning experiences should be developed to enable students to achieve the objectives? How should the learning experiences be organized to increase their cumulative effect? How should the effectiveness of the curriculum be evaluated?" (p.42). The evaluation applying Tyler's objective model has three basic steps. Firstly, the instructional objectives are specified. Secondly, the performance data is collected. The final step is to compare the performance data with the objectives specified.

2.2. Strengths and weaknesses of Tyler's objective model

Tyler's objective model has several strengths in evaluation, especially in curriculum development. Obviously, Tyler's model is based on the objective-oriented theory and this model considers curriculum as a means of aiming toward an educational object. The model aims at student's developing behaviors as the curriculum target of teaching. It is certainly one innovation in the field of education in the early 20th century as the application of the Tyler's model which is now used as an approach that helps to have better curriculum development in regard to the curriculum's objectives (Chen, Chen, & Cheng, 2005). In fact, with the model's development and innovation, Tyler's objective model is now used to define objectives for the new curricula and assess the degree to which the objectives are later realized. Curriculum is viewed as a set of broadly planned school-experiences designed and implemented to help students achieve specified behavioral outcomes (Madaus & Stufflebeam, 2000).

Secondly, since Tyler's model engages internal comparison of outcomes with objectives, evaluations using this model do not apply for costly and disruptive comparisons between experimental and control groups. The model calls for the measurement of behaviorally defined objectives, so an

emphasis is put on learning outcomes instead of organizational and teaching inputs. In this case, the model helps to avoid the subjectivity of the professional judgment or accreditation approach (Madaus & Stufflebeam, 2000).

However, although the model is highly appreciated in regard to curriculum development, some criticisms on it can be found in the field of educational evaluation (Chen C., Chen Y. & Cheng, 2005; Huang & Yang, 2004; Stufflebeam & Shinkfield, 1985). Huang and Yang (2004) criticize that the model does not provide feedback mechanism to tell stakeholders or evaluators on how to deal with improvements. It means that the evaluation applying Tyler's objective model has little use in improving a program and assessing the worth of a program. In addition, there is no connection between evaluation and organization. In the same way, Stufflebeam and Shinkfield (1985) point out some weaknesses of Tyler's model like placing the evaluator in technical role, and focusing mainly on objectives.

Furthermore, despite the benefits that behavioral objectives bring to curriculum design, evaluation using Tyler's objective model have some limitations. First, the evaluation focusing on behavior fails to evaluate objectives. Second, the objectives does not apply to all subjects or the design of a subject content (Huang & Yang, 2004). The objectives of the program or projects are not always stable and they can be changed to suit the context of the program's implementation. Therefore, Tyler's objective model cannot be applied to evaluate such programs or projects.

The final limitation of the Tyler's objective model lies in its objective-based nature. As the objectives can be changed at any time of the implementation of the program, the evaluation fails to evaluate the program with unstable objectives. Any educational program with unclearly defined objectives cannot

be evaluated using Tyler's objective model (Chen et al., 2005).

3. Stake's responsive model

Robert E. Stake created a system for carrying out evaluation in education in 1970s. (Popham, 1995). The model was then developed with the name Stake's responsive model (Stake, 1975, 1983). Stake's responsive model is the model that "sacrifices some precision in measurement, hopefully to increase the usefulness of findings to persons in and around the program" (Stake, 2011, p.8).

3.1. *The nature and characteristics of Stake's responsive model*

The evaluations is considered to be responsive "if it orients more directly to program activities than to program intents; responds to audience requirement for information; and if the different value-perspectives present are referred to in reporting the success and failure of the program" (Stake, 1975, p.14). The responsive evaluation puts an emphasis on the "concerns of the primary stakeholders, gathered through conversations with these parties on an ongoing basis during the evaluation" (Spiegel, Bruning, & Giddings, 1999, p.2).

In responsive model, the evaluator is a full, subjective partner in the educational program who is really highly involved and interactive. The evaluator's role is to provide an avenue for continued communication and feedback during the evaluation process (Stake, 1975).

According to Stake, there is no single true value to anything, but the value is in the eye of the beholder. It means that there may be many valid interpretations of the same events, based on a person's point of view, interest, and beliefs. The duty of the evaluator is collecting the views, the opinions of people in and around the program (Stake, 1983).

3.2. *Strengths and weaknesses of Stake's responsive model*

Responsive model has several advantages. First of all, in responsive evaluations, questions are allowed to emerge during the evaluation process rather than being formulated. The evaluation applying responsive model helps evaluators to acquire a rapid understanding of the program and to determine which issues and concerns are the most important to various groups of stakeholders. Secondly, the responsive evaluation uses content-rich information to describe the program in the way that is readily accessible to audiences (Stake, 1983; Hurteau & Nadeau, 1985). Furthermore, the responsive evaluation provides audiences with the chance to react to the evaluator's feedback and interact with the evaluator regarding their issues and concerns (Paolucci-Whitcomb, Bright, Carlson, & Meyers, 1987). In other words, the values and perspectives held by different audiences are explicitly recognized, which provides a context to examine different concerns. To sum up, Stake's responsive model is really successful in producing such evaluation accessible to a large variety of stakeholders.

However, besides some advantages, Stake's responsive model also has drawbacks. The first disadvantage is that the application of the model requires much time as the process of evaluation using the model takes a long time (Popham, 1995). Secondly, it is not easy to apply the model to evaluate educational programs if the evaluator is not an experienced one (Hurteau & Nadeau, 1985). The third disadvantage comes from the high level of interaction between the evaluator and stakeholders. With such high interaction, the role of the evaluator is ambiguous and in this case the evaluator "serves as a resource person rather than a researcher" (Popham, 1995, p. 3). Finally, the model is very flexible; as a result, it may be not easy to maintain the focus of the evaluation, which may result in a failure to answer specific questions (Hurteau & Nadeau, 1985).

4. Scriven's goal free model

Goal-free evaluation model was created in 1972 by Michael Scriven. At that time, for the more effective management practice of education, much money was invested in education. Such action raised the need for evaluating the educational projects which were funded by the government. As one person taking part in the evaluation of these projects, Scriven realized that the evaluations were influenced by the project's goals, which led to the low quality of evaluations. Therefore, he proposed the new model called 'goal-free evaluation' in program evaluation which is defined as "a model in which official or stated program goals and objectives are withheld or screened from the evaluator" (Youker & Ingraham, 2013, p.51).

The term "goal" used here is different from the term "objective". Goals are defined as "broad statements of program's purposes or expected outcomes, usually not specific enough to be measured and often concerning long-term rather than short-term expectations"(Weiss & Jacoss, 1988, p.528). By contrast, objectives are 'statements indicating the planned goals or outcomes of a program or intervention in specific and concrete terms' (Weiss & Jacoss, 1988, p.533).

4.1. The nature and characteristics of Scriven's goal-free model

As opposed to goal-based evaluation, goal-free evaluation puts a focus on the outcomes of the educational program, intended and unanticipated. The goal-free-evaluator does not deal with the rhetoric of the instructional designers regarding what they want to achieve, but rather pays attention to the results accomplished by the designers' educational programs (Popham, 1995).

The goal-free evaluation works in the way that it tries to discern what a total effects of the project are while assiduously avoiding the "false" information coming from the program or project goals. In this way, the side-effects

that may come from the goals of the projects or educational programs can be reduced. In other words, evaluators are not influenced by the goals of the projects or programs. However, it does not mean that the goal-free evaluation is recommended as the replacement of the goal-based evaluation. Goal-free evaluation is suggested to be used as a supplementary to goal-oriented framework by Scriven (Popham, 1995).

Youker and Ingraham (2013) suggested a guideline for evaluators to follow when they conduct a goal-free evaluation. The guideline proposes four main steps that evaluators should take as follows (p.7):

1. Identify relevant effects to examine without referencing goals and objectives
2. Identify what occurred without the prompting of goals and objectives
3. Determine if what occurred can logically be attributed to the program or intervention
4. Determine the degree to which the effect is positive, negative or neutral"

4.2. Strengths and weaknesses of Scriven's goal free model

One of the main benefits of the goal-free evaluation model is that it allows evaluators to be attentive to a wider range of program outcomes rather than just look for the program results that are stuck to the program aims/ goals. In this case, goal-free evaluators function as internal or external evaluators. For example, in a curriculum development project, one member of the project can be an internal evaluator who assess the worth of various project endeavors in terms of their results while another evaluator who is not the member of the project works as an external evaluator.

The second advantage of goal-free evaluation is that it can be used to supplement goal-based evaluation (Youker & Ingraham, 2013; Youker, Hunter, Bayer, & Zielinski, 2016). For instance, an evaluation may begin goal-free but later become goal-based using goal-free data for preliminary investigation

purposes while the evaluation is ensured to examine goal achievement (Stufflebeam & Shinkfield, 1985). In other words, the findings from GFE can be utilized as baseline information for subsequent GBEs. Moreover, a more comprehensive review can be accomplished when goal-free evaluation is used to supplement the findings from goal-based evaluation.

Another advantage of goal-free evaluation is avoiding the rhetoric true goal. “It is tragic when all resources go to goal-directed evaluation on a program when the stated goals do not even begin to include all important outcomes.”(Fitzpatrick, Sanders, & Worthen, 2004, p.85). Identifying which goals the evaluator should use is, in fact, a difficulty if the program has multiple stakeholders with different goals. GFE can avoid this issue by eliminating the distraction of goals (Youker & Ingraham, 2013).

Finally, goal-free evaluation can be adapted or adjusted to suit the sporadic changes in consumer needs, program resources, and program goals (Scriven, 1991; Davidson, 2005). Consumer needs, program foundation and environment are dynamic and may change over time; therefore, the goals of the program may not be relevant any more. In this case, the goal-free evaluator still continues with his task of recognizing and recording the effects providing that changes in goals or objectives are reflected in program’s actions and outcomes.

Nevertheless, even though goal-free model has a long history, it has remained conceptually abstract and highly theoretical with very few practitioners and others who have written about it (Youker & Ingraham, 2013). “Goal-free evaluation has been widely criticized for lack of operations by which to conduct it” (Shadish, Cook, & Levinton, 1991, p.61). In other words, it is quite hard for evaluators to assess educational programs using GFE as they just know the model in theory and there is a lack of knowledge of

the model in practice (Irvin, 1979; Mathison, 2005). The lack of knowledge in practice leads evaluators to the belief that they cannot use GFE in practice (Shadish, Cook & Levinton, 1991). More researches should be done in order to have a clear methodology and guidance for the GFE model to be effectively exploited in the world of evaluation.

5. CIPP model

Context, Input, Process, Product (CIPP) model is one of the oldest and thoroughly tested model which was developed by Daniel Stufflebeam in the late 1960s (Stufflebeam, 2000, 2003, 2014). The model was first created for helping improve and achieve accountability for the United States school programs. Later on, the model has been widely applied in many fields such as social programs, health professions, business, even in military and so on (Daniel L. Stufflebeam, 2000). The model is defined as “a comprehensive framework for guiding evaluations of programs, projects, personnel, products and evaluation system” (Stufflebeam, 2003, p.31).

5.1. *The nature and characteristics of CIPP model*

CIPP model has four different dimensions: Context evaluation, Input evaluation, Process evaluation and Product evaluation (Stufflebeam, 2000, 2003; Vo, 2017). Context evaluation deals with assessing needs, problems and opportunities within a defined environment. In put evaluation is used to evaluate competing strategies, the work plans, and the budgets for the strategies chosen to implement programs or projects. Process evaluation is used to monitor and assess activities carried out during programs or projects’ implementation. Product evaluation helps to identify and evaluate short term, long term, intended or unintended outcomes of programs or projects (Shtufflebeam, 2000, 2003, 2014; Vo, 2017a, 2017b).

The model is based on “learning by doing” (Stufflebeam, 2014, p.318). CIPP model is a continuous effort to identify and correct mistakes made in evaluation practice, and it is also the way to invent and test needed new procedures for more effective practices. The underlying theme of CIPP model is that the most important purpose of evaluation is “not to prove but to improve” (Stufflebeam, 2003, p.58).

CIPP model underlines both summative and formative evaluations. CIPP evaluations are formative when keying the collection and reporting information for improvements while they are considered to be summative when they assess the completed project or program activities or performances of services. With the summative role, the evaluations sum up the value meanings of relevant information

and put a focus on accountability (Stufflebeam, 2000, 2003). Besides, the model has objectivist orientation based on the theory that “moral good is objective and independent of personal or merely human feelings” (Stufflebeam, 2000, p.281). In other words, the evaluation applying CIPP model is free from human subjective feelings to reach a more precise conclusion.

Evaluator can use the whole CIPP model for their evaluation or just select one dimension of CIPP model for use. Stufflebeam (2000) proposes a specific guideline for evaluators, which states the objectives, methods and the use of each component of CIPP model. Based on the guideline, evaluators can decide which component or dimension of CIPP they should use for their evaluation, and what method they should adopt.

Table 1. Objectives, methods and relation to decision in the change process of four CIPP dimensions

	Objectives	Methods	Relation to decision making in the change process
Context	<ul style="list-style-type: none"> - To define the institutional/ service context - To identify the target population and assess its needs - To identify pertinent area assets and <i>resource opportunities</i> for addressing the needs - To diagnose <i>problems</i> underlying the needs - To judge whether goals are sufficiently responsive to the assessed needs 	<ul style="list-style-type: none"> - By using such methods as survey, document review, secondary data analysis, hearings, interviews, diagnostic tests, system analysis, and the Delphi technique 	<ul style="list-style-type: none"> - For deciding on the setting to be served - For defining <i>goals</i> and setting <i>priorities</i> - For surfacing and addressing potential <i>barriers</i> to success - For providing assessed needs as a <i>basis for judging outcomes</i>
Input evaluative	<ul style="list-style-type: none"> - To identify and assess system <i>capabilities</i> and alternative service <i>strategies</i> - To closely examine planned <i>procedures, budgets, and schedules</i> for implementing the chosen strategies 	<ul style="list-style-type: none"> - By <i>inventorying</i> and <i>analyzing</i> available human and material resources - By using such methods as <i>literature search, visits to exemplary programs, advocate teams, and pilot trials</i> to identify and examine promising solution strategies - By <i>critiquing</i> procedural designs for relevance, feasibility, cost and economy. 	<ul style="list-style-type: none"> - For selecting <i>sources of support</i> and solution <i>strategies</i> - For explicating a sound procedural <i>design</i>, including a budget, schedule, and staffing plan - For providing a basis for <i>monitoring and judging implementation</i>

<p>Process evaluative</p>	<ul style="list-style-type: none"> - To identify or predict <i>defects</i> in the procedural design or its implementation - To provide information for the programmed decisions - To record procedural events and activities for later analysis and judgment 	<ul style="list-style-type: none"> -By <i>monitoring</i> the activity’s potential procedural barriers and remaining alert to unanticipated ones - By obtaining specified <i>information for programmed decisions</i> - By <i>interviewing</i> beneficiaries, <i>describing</i> the actual process, maintaining a <i>photographic record</i> and continually <i>interacting</i> with and observing the activities of staff and beneficiaries 	<ul style="list-style-type: none"> - For <i>implementing and refining the program design and procedures</i>, i.e. for effecting process control - For lodging the actual process to provide a <i>basis for judging implementation and interpreting outcomes</i>
<p>Product evaluative</p>	<ul style="list-style-type: none"> - To collect descriptions and judgment of <i>outcomes</i> - To relate outcomes and judgment of <i>outcomes</i> - To relate outcomes to goals and to context, input and process information - To interpret the efforts’ merit and worth 	<ul style="list-style-type: none"> - By operationally defining and <i>measuring</i> outcomes - By <i>collecting judgments</i> of outcomes from stakeholders - By performing both <i>qualitative</i> and <i>quantitative</i> analyses - By <i>comparing outcomes</i> to assessed needs, goals, and other pertinent standards 	<ul style="list-style-type: none"> - For deciding to <i>continue, terminate, modify</i> or <i>refocus</i> a change activity - For presenting a clear <i>record of effects</i> (intended and unintended, positive and negative) - For <i>judging</i> the effort’s merit and worth

(Stufflebeam, 2000, p.302)

5.2. Strengths and weaknesses of CIPP model

CIPP model has a long history and it has been updated regularly, so it proves to be extremely beneficial in evaluation. First of all, the model is not designed for any specific programs or solution in mind; therefore, it is easily applied to multiple evaluation situations. It is used by “evaluators, program specialists, researchers, developers, policy groups, leaders, administrators, committees or task group, and layer persons” (Stufflebeam, 2014, p.310). Zhang from East Carolina University stated that while she searched for relevant literature on CIPP model, she found about 200 CIPP related evaluation studies, journal articles, and doctoral dissertations in many nations and in many fields (Stufflebeam, 2014). The model is also found to be applied in 134 doctoral dissertations at eighty- one universities involving 39 disciplines. CIPP model is employed in such disciplines as agriculture,

aviation, business, communication, distance education, elementary, tertiary, and secondary religion and sociology.

Secondly, the four different dimensions (Context, Input, Process, Product) of CIPP model can be used as the whole process to evaluate programs or projects or can be applied separately to suit the need of the evaluation. Context evaluation is used for planning decisions to determine objectives; Input evaluation helps to structure decisions to design instructional procedures; Process evaluation is to implement decisions to use, monitor and improve procedures; Product evaluation is applied for recycling decisions to judge and to react to the outcomes produced by procedures (Popham, 1995).

Thirdly, it is easy to apply CIPP model in evaluation as the model has a clear guidance developed by Daniel Stufflebeam and his colleagues. Stufflebeam (2000, 2003, 2014) provides evaluators with in-depth guidance

on when, why and how to use CIPP model. Based on such guidance, evaluators can decide whether they apply all CIPP model or just choose a particular dimension to use in their evaluation. A comprehensive framework of CIPP model is very useful for evaluators during their evaluation. Besides, Stufflebeam (2007) assists evaluators by a checklist that they can use to figure out what they need to do during the evaluation. The checklist also helps evaluators to know what they should deal with in their evaluations.

Despite its being widely used, CIPP has some drawbacks that should be considered before it is applied for evaluation. First of all, it is said to be similar to the needs assessment. Context evaluation has some features in common with needs assessment when it also deals with needs. Secondly, the application of the model takes much time if the whole model is applied.

6. Conclusion

Currently, evaluation in education is greatly paid attention when the quality of education and the improvement that needs to be made to educational programs are the main concern of the whole educational system. The application of the above discussed models is essentially useful for evaluators in Vietnam. Universities can make use of any of the above discussed models to evaluate their programs for improvement. These models are also useful for curriculum development of the new programs in Vietnam. Among the four models, the theory of Tyler's objective model is often applied in curriculum development or evaluation. Evaluators in Vietnam can apply Tyler's four principles to develop the program's curriculum. Besides, Tyler's objective model can be used to assess whether the program meets its intended objectives. Responsive model is used when the time for evaluation is limited. In case, program's designers want to have information for the

program's improvement quickly, responsive model is a good choice. Scriven's goal-free model is carried out as a supplement for goal-based model to increase the effectiveness of the evaluation. The final model, CIPP, is widely used for educational programs or projects in many fields for not only accountability but also improvements. The whole CIPP model is suitable for universities which are under accreditation because the model provides chances for evaluators to assess not only programs' implementation but also universities as the system.

All in all, the four models discussed above are well-known evaluation ones with long history. Each model has its own strengths and weaknesses when being applied for evaluation. Therefore, evaluators should consider their needs and also their evaluation experiences to select the one suitable for their evaluations.

References

- Chen, C., Chen, Y., & Cheng, K. (2005). *A study on comparing the objective model in curriculum planning between Taiwan and America*, 1–12. Retrieved from <http://rnd2.ncue.edu.tw/ezcafiles/b004/img/img/316/96-1-8p.pdf>
- Davidson, E. J. (2005). *Evaluation methodology basics: The nuts and bolts of sound evaluation*. Thousand Oaks, CA: Sage.
- Fitzpatrick, J., Sanders, J., & Worthen, B. (2004). *Program evaluation* (3rd editio). Boston: MA: Pearson.
- Huang, G. H., & Yang, L. L. (2004). *Curriculum development and design: Concept and practice*. Tapei: Hsi Ta Shu Yuan.
- Hurteau, M., & Nadeau, M. (1985). *The pros and cons of responsive model*.
- Irvin, J. (1979). Goal-free evaluation: Philosophical and ethical aspects of Michael Scriven's model. *The CEDR Quaterly*, 12(3), 11–14.
- Lynch, B. K. (1996). *Language program evaluation: Theory and practice*. New York: Cambridge University Press.
- Madaus, G. F., & Stufflebeam, D. L. (2000). Program evaluation: A historical overview. In D. L. Stufflebeam & G. F. Madaus (Eds.), *Evaluation models: Viewpoints on educational and human services evaluation* (Second edi, pp. 3–18). Kluwer Academic Publisher.

- Mathison, S. (2005). *Encyclopedia of Evaluation*. Thousand Oaks, CA: Sage.
- Paolucci-Whitcomb, Bright, W. E. I., Carlson, R. V., & Meyers, H. W. (1987). Interactive evaluation: Processes for improving special education leadership training. *Remedial and Special Education*, 8(3), 52–61.
- Popham, W. J. (1995). *Educational evaluation* (3rd editio). Boston: Allyn and Bacon.
- Scriven, M. (1974). Pros and cons about goal free evaluation. In W. J. Popham (Ed.), *Evaluation in education: Current applications* (pp. 34–67). McCutchan Publishing Corporation.
- Scriven, M. (1991). *Evaluation thesaurus* (4 edition). Newbury Park, CA: Sage.
- Shadish, Cook, W. R., & Levinton, L. C. (1991). *Foundation of program evaluation*. Newbury Park, CA: Sage.
- Spiegel, a M. Y. N., Bruning, R. H., & Giddings, L. (1999). Using Responsive Evaluation to Evaluate a Professional Conference. *American Evaluation Association*, 20(1), 57–67. <https://doi.org/10.1177/109821409902000105>
- Stake, R. E. (1975). *Evaluating the arts iin education: A responsive approach*. Columbus: Charles E. Merill Publishing Company.
- Stake, R. E. (1983). Responsive Evaluation. In T. Husen & T. N. Postlewaite (Eds.), *International encyclopeddia of education: Research and studies*. Newyork: Pergamon Press.
- Stake, R. E. (2011). Program evaluation particularly responsive evaluation. *Journal of Multidisciplinary Evaluation*, 7(15).
- Stufflebeam, D. L. (2000). The CIPP Model for Evaluation. In D. L. Stufflebeam, G. F. Madaus, & T. Kellaghan (Eds.), *Evaluation Models: Viewpoints on Educational and Human Services Evaluation* (second edi, pp. 279–318). Boston: Kluwer Academic.
- Stufflebeam, D. L. (2003). The CIPP Model for Evaluation. In T. Kellaghan, D. L. Stufflebeam, & L. A. Wingate (Eds.), *International Handbook of Educational Evaluation* (pp. 31–62). Dordrecht: Kluwer Academic Publisher.
- Stufflebeam, D. L. (2007). Cipp evaluation model checklist. *NYLC Resources*, (March), 15. Retrieved from https://www.wmich.edu/sites/default/files/attachments/u350/2014/cippchecklist_mar07.pdf
- Stufflebeam, D. L. (2014). Daniel Stufflebeam’s CIPP Model for Evaluation- An Improvement and Accountability-Oriented Aproach. In D. L. Stufflebeam, Coryn, & L. S. Chris (Eds.), *Research Methods for Social Sciences: Evaluation Theory, Models and Applications* (pp. 310–339). Someset: Jossay-Bass.
- Stufflebeam, D. L., Coryn, & Chris, L. S. (2014). *Research method for social sciences: Evaluation Theory, Models and Applications*. Jossay-Bass.
- Stufflebeam, D. L., & Shinkfield, A., J. (1985). *Systematic evaluation: A self-instructional guide to theory and practice*. Boston: Kluwer Academic Publisher.
- Tyler, R. W. (1949). *Basic principles of curriculum and instruction*. Chicago: The University of Chicago Press.
- Tyler, R. W. (1976). Two new emphases in curriculum development. *Educational Leadership*, 34, 61.
- Vo, T. K. A. (2017a). Evaluating the implementation of action research course in an in-service teacher training program in Vietnam. *Journal of Nusantara Studies*, 2(2), 88–97.
- Vo, T. K. A. (2017b). Stufflebeam’s Context, Input, Process and Product evaluation model: An improvement, accountability approach for evaluation in education. In *8th international conference on TESOL*. Hochiminh: SEMEO RETRACT.
- Weiss, H. B., & Jacoss, F. H. (1988). *Evaluating family programs*. New York: Aldine de Gruyter.
- Youker, B. W., Hunter, O. C., Bayer, N., & Zielinski, A. (2016). Who needs goals? A case study of goal-free evaluation, *Journal of Multidisciplinary Evaluation*, 12(27), 27–43.
- Youker, B. W., & Ingraham, A. (2013). Goal-Free Evaluation: An Orientation for Foundations’ Evaluations. *The Foundation Review*, 5(4). <https://doi.org/10.9707/1944-5660.1182>

MÔ HÌNH ĐÁNH GIÁ CHƯƠNG TRÌNH GIÁO DỤC: ƯU VÀ NHƯỢC ĐIỂM

Võ Thị Kim Anh

Nghiên cứu sinh, Universiti Malaysia Sabah

*Khoa Tiếng Anh, Trường Đại học Ngoại ngữ, Đại học Đà Nẵng,
131 Lương Nhữ Hộc, Khuê Trung, Cẩm Lệ, Đà Nẵng, Việt Nam*

Tóm tắt: Đánh giá giáo dục đã được chú trọng và các mô hình đánh giá giáo dục được đề xuất trong thế kỷ 20 đã được phát triển và áp dụng rộng rãi trong đánh giá giáo dục ở thế kỷ 21. Bài báo cung cấp cho độc giả cái nhìn tổng thể về bốn mô hình đánh giá nổi tiếng trong lĩnh vực giáo dục: mô hình khách quan của Tyler, mô hình phản hồi của Stake, mô hình phi mục tiêu của Scriven và mô hình CIPP của Stufflebeam. Những mô hình này có một bề dày lịch sử và được phát triển theo thời gian. Bốn mô hình này được áp dụng rộng rãi trong nhiều lĩnh vực đánh giá, nhưng chủ yếu là trong đánh giá chương trình giáo dục. Với mục đích giúp các nhà đánh giá giáo dục hiểu rõ hơn về các mô hình này, bài báo trình bày bản chất của các mô hình, đặc điểm của mô hình, cũng như thảo luận điểm mạnh và điểm yếu của từng mô hình.

Từ khóa: đánh giá giáo dục, mô hình đánh giá, mục tiêu