



Approaches to Evaluation: in Pursuit of Appropriacy For EFL Pre-Service Teacher Training Programs in Indonesia

Didin Nuruddin Hidayat

Department of English Education, UIN Syarif Hidayatullah Jakarta, Indonesia

didin.nuruddin@uinjkt.ac.id

ABSTRACT

Keyword:

Approaches to Evaluation, EFL Teaching in Indonesia, Program Quality.

Evaluation has the potential to support the improvement of education as well as its specific programs, products, and materials. Also, evaluation can uncover the relevance, effectiveness, and efficiency of an education program. Another potential role of evaluation is that it can identify areas for improvement, which further can help develop a more efficient program in the future. The present study seeks to investigate the most appropriate approach to evaluation for the EFL pre-service teacher training programs in Indonesia. Numerous potential approaches were identified and discussed: Pseudoevaluations, Questions Methods-Oriented, Improvement Accountability, and Social Agenda/Advocacy. The present study argues that the Improvement/Accountability Approaches are the most suitable evaluation approach to apply to the pre-service EFL teacher training programs in Indonesia, due to its strength in the philosophical stance that fits with the Indonesian context. Also, the purpose of the evaluation and the various types of informants offer by the approach add to the appropriacy. The study may provide significant contributions to better design more effective teacher training programs, which will lead to producing more qualified graduates.



Article History:

Received: 05-10-2018
Revised : 27-12-2018
Accepted: 28-12-2018
Online : 30-12-2018



This is an open access article under the **CC-BY-SA** license



<https://doi.org/10.31764/ijeca.v1i3.2129>

A. INTRODUCTION

Evaluation plays an important role in a program. One of the notable roles of evaluation is that it can be utilized to assess the success or failure of a program. Evaluation can also provide detailed descriptions of a program’s achievements and strengths. Another potential role of evaluation is that it can identify areas for improvement, which further can help develop a more efficient program in the future. Evaluation is essential in enhancing and ensuring the quality improvement of education. Sceeren et al., (2003) argued that there are at least three primary reasons for the need for evaluation in a program. Evaluation is needed to “... formally regulate desired levels of quality of educational outcomes and provisions; to hold educational service providers accountable and to support ongoing improvement in education (p. 3)”. The potential

benefit of evaluation is also echoed by Hill et al., (2013) where they elucidated that evaluation may improve education quality since evaluation can uncover the relevance, effectiveness, and efficiency of an education program.

In light of these pivotal roles of evaluation, in the past few decades, scholars have paid growing concern towards evaluation through seminal writings on evaluation and identification of approaches to evaluation which at the end lead to the development of program evaluation models (e.g., Sceeren et al., 2003; Scriven & Coryn, 2008; Stufflebeam & Shinkfield, 2007). While extensive evaluation models can be found in the literature, limited evaluation models designed explicitly for EFL pre-service teacher training programs can be found, especially in the Indonesian context.

Although the explosion in the interest of language program evaluation just started in the 1990s, from then on, scholars have shown eagerness to explore the area and have started to produce extensive pieces of research studies in different countries. A study by Stewart (2007) in Japan, employing both teacher and student participants, revealed that the ways teachers perceive the learning tasks tend to be different from how the students think. It cannot be denied that both teachers and learners have a strong influence on the outcome of a program. Thus, incorporating multiple angles on approaching tasks leads to numerous benefits. Stewart's (2007) evaluation of his self-taught language program found ways of narrowing the distance of the perceptions of teachers and students regarding the learning tasks. Luchini (2002) did an evaluative case study in Argentina concerning the teachers' methodology of teaching writing. It was found that teachers responded positively to the combined methodology and practice offered on the project. Likewise, students mean scores in writing increase quite significantly. The study also shed some potential insights on how the writing project can be more effective in its subsequent implementations.

Klassen et al., (1997) evaluated an English program in Hongkong. It attempted to reveal the students' perception of the program effectiveness. The study uncovered that students perceive the self-access mode as of assistance to trigger their motivation. The positive learning attitudes of students through the self-access mode were also noticeable as it enhanced their four language skills and improved their confidence. The findings of the study then suggest ways of how the program can be improved by a design that consists of three interrelated-concept: directed, semi-directed and self-directed stages of learning. Using a mixed-method design and gaining the perspectives from students and teachers, Chang's (2010) evaluation study in Taiwan higher education English institution revealed that even though the learners were not perceived as having a good level of comprehension of the lectures delivered in English, they believe that the English instruction supports the improvement of their English proficiency, more specifically in the skills of listening. Further, the findings suggested two recommendations for the university. The first is offering voluntary EAP language courses, and the second one is restricting the EMI only to be joined by students having sufficient English proficiency.

Considering the above-mentioned conditions, the present study aims to explore the appropriate evaluation design that fits in the Indonesian context, more particularly in

the tertiary education level. This study strives to contribute to the knowledge base by filling the gap of the limited availability of specific evaluation models designed for EFL pre-service teacher training programs. The model would be beneficial in assisting language teacher educators in evaluating EFL teacher training programs.

B. METHODS

The present study employed a literature web-based survey research. There are two kinds of methods of web-based surveys: non-probability and probability (Couper, 2000). The present study adopted the latter method, probability, by finding the sample from the populations. The operational technique was conducted through a quick Google search on the internet. After conducting a quick search through the internet, it was found some promising evaluation models that have been used by language researchers and educators in evaluating the effectiveness of English language and education programs.

The search for evaluation models from the internet that fit for English language and education programs generated several results. Among others are the Context-Input-Process-Product evaluation model developed by Stufflebeam (1983). A further widely-used evaluation model for education programs found on the internet was the four-level evaluation model developed by Kirkpatrick and Kirkpatrick (1996), which consisted of four levels: reaction, learning, behavior, and results. Another promising evaluation model found from internet search and was applied widely by language researchers and educators was an evaluation model pioneered by Peacock (2009). Concerning a more updated evaluation model and also a closer look at the evaluation program focusing on improving the quality of English training programs, the researcher chose Peacock (2009) as one of the promising evaluation models that may cater best to the needs of English training programs in the Indonesian context.

C. RESULT AND DISCUSSION

1. Studies On English Teacher Training Evaluation

More evaluation studies, especially on English teacher training evaluation, have enriched the discussion of evaluation. Yavuz and Topkaya (2013) conducted an evaluation study of the Turkish Higher Education Council's (HEC) English Language Teacher Education Program (ELTEP). The main changes fall into five categories: removed courses, added courses, re-sequenced courses, class hour changes, and modified courses. The study investigated the teachers' perspectives on the new changes by employing a qualitative design using open-ended questionnaires. The study revealed that even though the added courses were found appropriate, the other categories received heavy criticism from the teachers. Further, the participants shared the same views that the top-down approach employed by the council in restructuring the teacher education program did not take into account the teachers' opinions, experiences, and practices. Also, there was evidence of a lack of coordination among the faculties, Higher Education Council, and Ministry of National Education in formulating new teacher training programs, which thus resulted in minimum cooperation and coordination.

Liton (2013) presented the research study of English for Specific Purposes (ESP) teachers training. The study attempted to uncover the training program effectiveness drawn from the perspectives of 46 ESP teacher trainees joining the training program. The study found that the effectiveness of the program is evident in enhancing the teachers' teaching

performance. After completing the training program, the teacher trainees demonstrated better performance in real classroom practices. Besides, the program was also believed to cater to the students' pragmatic aspects and demands in the workplace environment. The study further promoted potential suggestions for the future training program, including the availability of similar ESP teachers training programs as it is essential to uplift their professional teaching capabilities.

Coşkun and Daloğlu (2010) investigated an English language teacher training in Turkey. The study seeks to investigate the program effectiveness drawn from multiple stakeholders' perspectives. Gaining perspectives from 55 final year students and three English instructors, the study revealed that though participants share the same views on some components of the programs, the different view lies in the way they perceive the balance between linguistic and pedagogic competences. The study then proposes some recommendations for the future teacher training program that consists of the inclusion of training in language teaching methodologies, a framework of teaching practice, and a course in language proficiency.

2. Studies On English Teacher Training Evaluation In Indonesia

Though the discussion of an evaluation study in Indonesia is not as productive as in other countries, the evidence is still worth discussing. Putra and Sitawati's (2012) study aims at investigating whether the English curriculum of the Business Administration Department of Politeknik Negeri Bali is relevant to the industry needs. Gathering data using questionnaires from 75 student participants, it was revealed that, in general, students perceive that the English curriculum meets the needs of the industry. Though there was no specific evaluation model applied in the study, the researchers developed their data collection instrument in the form of a questionnaire with a 5-point Likert scale. The study suggests that with the dynamic nature of the industry, it is essential to consistently review and make the curriculum compatible with any dynamics occurring in the industry.

Attempting to seek ways of quality improvement of English training programs, Suhuri (2008) conducted an evaluative study on the Indonesian high school context. He developed an evaluation model drawing from the work of Kirkpatrick. There are two core components in the evaluation of the learning and teaching model, named as *Evaluasi Pembelajaran Bahasa Inggris (EPBI)*. The first component assesses the teaching-learning process covering class teacher performance, teacher personality, student behavior to learning and teaching English, as well as English teaching-learning facilities. The second core component assesses the teaching-learning outcomes covering the four language skills of English. Suhuri (2008) claimed that his model of evaluation could be regarded as a good model as it has been empirically tested. He further stated that the model has comprehensive coverage, practical guidance, valid data collection instruments, and a concordant model with the data field.

There is scant on information on the evaluation of English teacher education programs in the Indonesian context. It indicates that little work of this kind, focusing specifically on the evaluation of ELT teacher training programs, has been conducted. The studies might be available, but perhaps are only intended for internal consumptions and are not disseminated to the public domain.

3. The Need For Program Evaluation

Brown (1989) asserted that evaluation is "the systematic collection and analysis of all relevant information necessary to promote the improvement of the curriculum, and assess its effectiveness and efficiency, as well as the participants' attitudes within a context of particular institutions involved (p. 223)". This definition implies that a need for systematic

evaluation in a program is crucial in terms of improving the quality of a program. It is conducted by finding out how a program is implemented and put into practice. It can also reveal the strengths and weaknesses of a program so that informed judgment on what aspects of a program that must be kept, modified, added, or deleted can be applied. Another benefit is assessing whether an implemented program has met its objectives. Evaluation measures whether the objectives set earlier have been achieved. Evaluation can also reveal if the program has well served the stakeholders. Lynch (1996) highlighted two benefits of evaluation in that it does not only offer valuable information to insiders on how effective the current program is going and which aspects can be improved, but also provide insights to outsiders on the program's accountability.

4. Approaches To Evaluation

The increasing number of alternative approaches to evaluations has started in the middle of the 1970s. Scholars have intensively developed approaches to evaluation and classified them based on, for instance, philosophy (Guba, 1969), ideology (House, 1978), and orientation (Stufflebeam, 2002). An orientation-based classification is grouped into four significant approaches comprising a total of 22 approaches: Pseudoevaluations (2 approaches), Questions/Methods-Oriented (13 approaches), Improvement/Accountability (3 approaches), and Social Agenda/Advocacy (4 approaches).

Pseudoevaluations approaches cannot be accepted as a sound evaluation practice due to their nature that misrepresent value interpretations of some objects or programs. Likewise, pseudo evaluation approaches also tend to promote invalid or incomplete findings (Stufflebeam, 2002). The study included in Pseudoevaluations is the politically oriented studies where the released result of evaluation is only the one that meets clients' expectations. Another unfavorable result tends to be disregarded and is kept confidential. Public relations studies also belong to the category of Pseudoevaluations. These studies attempt to create positive images of objects or programs without considering the actual occurring condition. The results of evaluations seem to be shaded, or sometimes false judgments are made to build the positive images.

Questions/Methods-Oriented approaches address specified questions and employ a particular method. These approaches can also be regarded as quasi-evaluation studies as they provide evidence of a full evaluation of the strengths and weaknesses of a program, though they often have narrow evaluation focus. Stufflebeam (2002) criticized for these approaches "...emphasize technical quality and posit that it is usually better to answer a few pointed questions well than to attempt a broad assessment of something's merit and worth (p. 40)". Some of the approaches that belong to Questions/Methods-Oriented are objective-based studies, performance testing, experimental studies, program theory-based evaluations.

Improvement/Accountability Approaches are grounded on the premise that comprehensively assessing the worth and merit of a program is a focal need. This is conducted by developing a complete set of criteria and questions in assessing the value of a program. The approaches gather relevant information from stakeholders often by employing a mixture of qualitative and quantitative approaches in triangulating the findings. The three approaches that belong to the improvement/accountability approaches are Decisions/Accountability, Consumer-Oriented, and Accreditation. Stufflebeam (2002) claimed that "T[t]he three approaches emphasize respectively improvement through serving program decisions, providing consumers with assessments of optional programs and

services, and helping consumers to gain assurances that given programs are professionally sound and effective (p. 68)".

Social Agenda/Advocacy approaches aim to promote a better quality in society from an evaluation of a program empowering the disadvantaged in society. The approaches believe that all segments in society deserve to receive the same treatment as education and social opportunities. There are four approaches that belong to this group: Client-Centered Studies, Constructivist Evaluation, Deliberate Democratic Evaluation, and Utilization-Focused Evaluation. All approaches "... are oriented to employing the perspectives of stakeholders as well as of experts in characterizing, investigating, and judging programs. They favor a constructivist orientation and the use of qualitative methods" (Stufflebeam, 2002, p. 68). The main weakness of these approaches is that they are unlikely to successfully meet the standard of a proper evaluation as these approaches put much attention to the social mission.

Considering the condition of the English teacher training program in Indonesia, it seems that the Improvement/Accountability Approaches are the most suitable evaluation approach for the Indonesian context. All approaches described above have their inherent strengths and weaknesses. It is the evaluator's own decision on which approach to be utilized as a philosophical stance. Ideally, when selecting an appropriate approach for an evaluation, one should consider the purpose of the evaluation and the type of information needed to meet the purpose. There is no single evaluation approach can be claimed as the best evaluation approach for different contexts of evaluation. However, Stufflebeam (2002) rated the value of all approaches above based on the standard evaluation categories consisting of utility, feasibility, propriety, and accuracy, with rating scores from excellent to poor. The most reliable and most promising approach for continued use and development is Improvement/Accountability (Decision/Accountability Approach), followed by Questions/Method (Case Study Approach), and Social Mission/Advocacy (Utilization/Focused Approach) respectively. Meanwhile, Pseudoevaluations approaches are not ranked as they fail to demonstrate rigor and relevance.

Another justification for selecting the Improvement/Accountability Approaches is based on the grounds that this approach offers a holistic approach in evaluating the worth and merit of the English teacher training program. Improving the condition of English teacher training in Indonesia is necessary as it leads to quality improvement. Further, the Improvement/Accountability Approaches acknowledge the contribution of various stakeholders involved in the teaching-learning process. It allows the opportunity to obtain rich insights from different perspectives, e.g., teachers, students, coordinators, school principals, heads of department, and so forth. By collecting their perspectives, it is hoped that a complete picture of an evaluation can be obtained and thus be able to take better actions from the results of the evaluation. This, in turn, ultimately allow more significant opportunities for program quality improvement.

D. CONCLUSION AND SUGGESTIONS

The study has underlined an evaluation model for EFL pre-service teacher training programs at the tertiary level in Indonesia. The model is intended to better assist educators in evaluating the existing programs. The need for evaluation and an established model of teacher training evaluation is vital, considering research studies

findings revealing the benefits of evaluation. The study may also provide a significant contribution to better design more effective teacher training programs, which will lead to producing more qualified graduates whom most of them will teach at primary, secondary and even higher education levels. The study also expands and contributes to the plethora of discussion in the English teacher training evaluation. Compared to the studies on mainstream education, research studies on language program evaluation is still in the infancy, more particularly in the context of Indonesia. Scholars called for more rigorous studies on evaluation considering the potential benefits it has in providing guidelines for evaluators in improving the future evaluation.

REFERENCES

- Brown, J. D. (1989). Language program evaluation: A synthesis of existing possibilities. In R. K. Johnson (Ed.), *The Second Language Curriculum* (pp. 222–241). Cambridge University Press.
- Chang, Y. (2010). English-medium instruction for subject courses in tertiary education: Reactions from Taiwanese undergraduate students. *Taiwan International ESP Journal*, 2(1), 55–84.
- Coşkun, A., & Daloğlu, A. (2010). Evaluating an English Language Teacher Education Program Through Peacock's Model. *Australian Journal of Teacher Education*, 35(6), 24–42.
- Guba, E. G. (1969). The failure of educational evaluation. *Educational Technology*, 9(5), 29–38.
- Hill, H. C., Beisiegel, M., & Jacob, R. (2013). Professional development research: Consensus, crossroads, and challenges. *Educational Researcher*, 42(9), 476–487.
- House, E. R. (1978). Assumptions underlying evaluation models. *Educational Researcher*, 7(3), 4–12.
- Klassen, J., Detaramani, C., Lai, E., Patri, M., & Wu, J. (1997). *Evaluating the English Foundation Programme in the self-access mode at City University. Case Studies of Improving and Learning from the Action Learning Project*.
- Liton, H. A. (2013). An assessment of the efficacy of Engineering ESP teachers' training program in promoting better performance at ESP classroom. *English for Specific Purposes World*, 37(13), 1–17.
- Luchini, P. L. (2002). Evaluation of teachers' methodology of writing skill teaching: A case study. *The Linguistic Association of Korea Journal*, 10(4), 211–232.
- Lynch, B. K. (1996). *Language program evaluation: Theory and practice*. Cambridge University Press.
- Putra, K. D. C., & Sitawati, A. A. R. (2012). Evaluasi kurikulum bahasa Inggris jurusan Administrasi Niaga Politeknik Negeri Bali. *Stilistetika Journal*, 1(1), 28–45.
- Sceeren, J., Glas, C., & Thomas, S. M. (2003). *Educational evaluation, assessment, and monitoring: A systemic approach*. Swets and Zeitlinger Publishers.
- Scriven, M., & Coryn, C. L. S. (2008). The logic of research evaluation. *New Directions for Evaluation*, 118, 89–105.
- Stewart, T. (2007). Teachers and learners evaluating course tasks together. *ELT Journal*, 61(3), 256–266.
- Stufflebeam, D. L. (2002). Foundational models for 21st century program evaluation. In D. L. Stufflebeam, G. F. Madaus, & T. Kellaghan (Eds.), *Evaluation models: Viewpoints on educational and human services evaluation* (pp. 33–84). Kluwer Academic Publishers.
- Stufflebeam, D. L., & Shinkfield, A. J. (2007). *Evaluation theory, models, and applications*. John Wiley & Sons.
- Suhuri, S. (2008). *Model evaluasi pembelajaran Bahasa Inggris SMA*. Universitas Negeri Yogyakarta.
- Yavuz, A., & Topkaya, E. Z. (2013). Teacher educators' evaluation of the English language teaching program: A Turkish case. *Novitas-ROYAL (Research on Youth and Language)*, 7(1), 64–83.