

DEVISING AN EVALUATION APPROACH FOR A MALAYSIAN ARMED FORCES ENGLISH LANGUAGE PROGRAM

Muhammad Salim Tufail*, Hilmi Azani Husain, Dinie Asyraf Salehan and Muhammad Shafiq Azid

Language Centre, National Defence University of Malaysia, Sungai Besi Camp, 57000 Kuala Lumpur, Malaysia

*Email: salim@upnm.edu.my

ABSTRACT: Program evaluation is becoming an important area in the development and improvement of language programs for the benefit of all stakeholders concerned. The progress in the field of language program evaluation has led to a greater degree of its understanding among the language teaching sphere. This paper looks at the author's experience in developing an evaluation approach for a Malaysian Armed Forces English language program. The approach was based on the works of some leading experts in the field. It carries out a critical analysis of the dimensions of evaluation (the summative-formative, product-process, and quantitative-qualitative dimensions) in the process of devising the most comprehensive and inclusive approach possible. In line with the opinions of the relevant leading experts, the author's experience found that the opposing elements of the dimensions must not be considered as dichotomies. Instead, they can be tailored to complement each other to suit a particular evaluation exercise based on the purpose, focal points and the type of information required. The outcomes of the study showed that an evaluation should utilize both perspectives by capitalizing on the strengths of each dimension to enhance the value of the measurable outcomes of the practice.

Keywords: language program, evaluation, approach, dimensions of evaluation, methodology

1. INTRODUCTION

The increasing importance of language program evaluation has led to a greater degree of its awareness within the language teaching fraternity and stakeholders alike, particularly to those acting as evaluators. Among the many challenges they face are to develop the most suitable, efficient and effective of evaluation approaches or models in carrying out their tasks. In this case, the experience of the author in carrying out an evaluation of a Malaysian Armed Forces English language program by expounding on the work of a leading expert [1] and supported by many others in seeking to uncover the underlying theoretical foundations of the evaluation process is presented. This understanding will then enable a potential evaluator to devise an approach which can be best tailored to evaluate a language program. The approach must be eclectic enough to unravel the many aspects of a language program which will serve to provide highly informed answers and solutions for the benefit of language teaching and program improvement. This paper thus takes a critical and in-depth look at the author's experience in devising an approach in the evaluation of a Malaysian Armed Forces English language program for officers [2]. In arriving at the evaluation approach, certain aspects and related factors were taken into account by the nature of their theoretical importance in the process. This approach was developed based on two main foundations. First, it looked into the nature of the evaluation process in terms of its purposes and focal points. Then, the 'dimensions of evaluation' [1] were weighed in to determine the approach and methodology. This knowledge was then applied in developing the approach based on the key components and processes of the language program.

2. THE IMPORTANCE OF ENGLISH LANGUAGE PROGRAMS IN THE MALAYSIAN ARMED FORCES

Today, the command of the English language among global citizens is more crucial than ever. This scenario is expected to be so in the future. In line with this development, the improvement of the English language proficiency within the military is becoming increasingly important in the professional development of its officers and men, particularly

in the midst of bustling globalization and the rapid progress in technology. Regardless of the situation, a good command of the English language most often serves to facilitate communications more efficiently and effectively not only in the more demanding day-to-day work but also in the ever challenging and complex military operations. The highly globalized and technologically sophisticated contexts of today are among the major factors in this phenomenon. In addition to this, a good command of the English language among military personnel also acts as an agent in the transfer of technology, thus serving as a catalyst of progress in the defense sector. It thus becomes crucial that military officers receive the best possible English language teaching and learning for the improvement of their proficiency so as to function with greater efficiency and effectiveness.

In this regard, language programs in the Malaysian Armed Forces need to be continuously improved along with the rapidly changing environment to meet more challenging and demanding needs. As mentioned by Richards [3], there is a need to explore factors in the language teaching context that are crucial in determining the success or failure of language programs. In this regard, evaluation of language programs is the process which can contribute to this need. Having said that, evaluation is a process which should not only provide an understanding of the complex workings of language programs but more importantly to contribute to their continuous improvement and value for money [1, 4, 5, 6]. This is why it is important that language programs are evaluated and in doing so, the most comprehensive, appropriate and effective approaches are devised for the maximum benefit of all parties concerned.

3. THE DEFINITION OF LANGUAGE PROGRAM EVALUATION

There is no single definition of language program evaluation as this field has evolved rapidly and encompassed a variety of purposes and methods. As such, many definitions have been written on this subject. Having said that, most of the definitions imply certain commonalities such as systematic efforts to define criteria, the need to obtain accurate information about program characteristics, and contribution

towards program improvement. For the purpose of this paper, the following definition by Brown [1] (p.222) will be used as a guide:

“...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”.

This definition is used as it represents practicality and comprehensiveness. At the same time, it concisely incorporates most of the key elements (within the curriculum) required in the conduct of language program evaluations.

4. DEVELOPMENTS IN LANGUAGE PROGRAM EVALUATION

Much progress has been made in the field of language program evaluation. Educational evaluation, from the early works by Ralph Tyler more than half a century ago, which primarily focused on the evaluation of student's achievement data. It was a product-oriented practice which was predominantly concerned with course outcomes, particularly test results, and is sometimes also referred to as assessment data [4]. However, since then the practice of educational evaluation has evolved significantly along with the changing environment, particularly with the advent of greater competition and decreasing public funds. This has led to a greater degree of stakeholder demands and the need to continuously improve language programs. Since then, the definition and understanding of the term 'evaluation' have evolved significantly along with the shifting educational landscape to gradually embrace its changing conceptualizations, scope, and purposes [4, 5, 7].

Today, the educational environment has taken a vastly different look. Educational practitioners and stakeholders now live in an age of growing accountability and transparency amid tighter budget constraints. This is exacerbated with the ever-increasing demands of stakeholders, including the public at large, practitioners and learners alike. These forces have led experts to seek more comprehensive, efficient and effective means to understand more about how language programs work (or fail) in the collective quest for providing support and solutions for continuous improvement and value for money where the enhancement of student learning is fundamental [5, 6, 8].

In other words, evaluation is not just answering how well students have done, but also addressing the wider-ranging questions. These include aspects such as how well the program has served the learners, educators and stakeholders, how much value for money the program has delivered, and how the program has fared in comparison to others or how effectively it has been executed in meeting its more eclectic objectives [4, 5]. With such demanding goals and pressures, the practice of language program evaluation has begun to see the need to better understand its theoretical foundations to effectively incorporate multi-faceted approaches into its practice. This understanding will help efforts to encapsulate the significant progress made in the field of language teaching and language program evaluation over the last half a century or so. It has now reached a stage where synthesizing methodologies and procedures in relation to the need to uncover the many elements that constitute the complexity of

language programs has become imperative. This is done to arrive at an approach best suited to answer wider-ranging key questions and stakeholder concerns [5, 6, 9, 10].

5. DEVISING AN EVALUATION APPROACH: SELECTING THE PURPOSE

There are many purposes of evaluation. It is one of the first aspects to consider before embarking on one. Evaluations can be carried out for a variety of reasons. One of the most important questions to be answered at the early stage is, “why is this evaluation required?” This question is important as it will have a bearing on the evaluation approach that is to be employed later. Broadly, two of the most common purposes of evaluations are as follows [6]:

a. **Formative evaluation.** It is carried out as part of the process of program development in order to discover out what is working well, what is not, and what problems need to be addressed. It focuses on on-going development and improvement of the program.

b. **Summative evaluation.** This type of evaluation seeks to make decisions about the worth or value of different aspects of the curriculum. Summative evaluation is concerned with determining the effectiveness of a program, its efficiency and to some extent, its acceptability. It takes place after a program has been implemented.

Both formative and summative evaluations were selected in the author's study as its objectives were to determine the program effectiveness, as well as to contribute to program improvement.

6. DEVISING AN EVALUATION APPROACH: SELECTING THE CHOICE OF FOCUS

The first step is to give the appropriate emphasis on the areas or focal points of the program to be evaluated. As a language program is a complex of many intricately linked components, its evaluation must have its own focal points based on what it is designed to inform [6, 8]. Among the focal points which can be selected are curriculum design, syllabus and program content, classroom processes, materials of instruction, the teachers, teachers' training, the learners, the institution and the learning environment.

The author focused on three major components of the language programme; (1) the teachers, (2) the learners, and (3) the teaching materials. These focal points were selected because these three aspects are not only fundamental issues in ELT [11], but also the most essential components in a language program. In addition, teachers and learners also play important roles in evaluation, both as stakeholders as well as key participants in the process [8, 10, 12]. Teaching materials, meanwhile, are just as important in a language program as they provide the corpus of the curriculum [10]. As mentioned by Lynch [14], teaching materials are one of the major determining factors of what gets taught in a language program.

In this case, the teachers, learners, and teaching materials were viewed as the core components of a language program [3, 6, 15]. By doing this, the author also fully acknowledged the importance of the teaching and learning element as the central process to which evaluation contributes to Mitchell [16], rather than the other way round. A deep understanding of these three intricately linked factors will shed light not just

on the end results, but more importantly the processes, i.e. what occurs in the program.

7. DEVISING AN EVALUATION APPROACH: THE FINAL STEP

The final step in the process is to develop an evaluation approach. This is the most challenging yet interesting stage as this is the point where the purpose and the focal points are linked to incorporate the most effective possible methodology into an approach which can allow this to happen. In order to envisage this complex process, it becomes imperative to weigh in what [1] calls as the ‘dimensions’ of evaluation. These dimensions can be viewed as the invisible or intangible ‘building blocks’ (if one may call it) of the evaluation process. They consist of what are generally regarded by researchers as opposing elements within particular paradigms, as follows:

- a. The summative-formative dimensions.
- b. The product-process dimensions.
- c. The quantitative-qualitative dimensions.

In program evaluations, the distinction between the opposing dimensions is based on differences in the type of information that can be yielded [1]. This will have an important bearing on the types of methodologies and procedures to be used.

The Summative-Formative Dimensions

Formative evaluation is carried out during the program implementation to find what is working well, and what is not, and what steps can be taken to improve the program [6, 8, 17]. Summative evaluation, meanwhile, is a terminal evaluation of a program that is already operational. Its purpose is to make judgments about a program's worth, its end result, or its effectiveness [1, 6, 17].

The Product-Process Dimensions

Next, the product and process dimensions are considered. The product dimension refers to the determination of whether the goals and instructional objectives of the program have been achieved [1], or the degree to which they are achieved [18]. Meanwhile, [1] and [18] said that the process dimension focuses on what was going on in a program (process), or its implementation, and this helps to arrive at the stated goals (product).

The Quantitative-Qualitative Dimensions

According to Creswell [19], qualitative research is a type of inquiry where its approach is useful for exploring and obtaining an in-depth understanding of a central phenomenon. The inquirer would ask respondents broad and general questions. The researcher then collects the in-depth views of the respondents and analyses the information for description and themes. Quantitative research, meanwhile, involves systematic empirical observations using statistical, mathematical or computational procedures. It is useful for describing trends and explaining the relationship among variables. According to Creswell [19], to conduct this inquiry, the investigator specifies narrow questions and locates or develops instruments to gather numerical data. The researcher then carries out statistical, mathematical or computational analysis of data to answer the research questions.

The descriptions of the different dimensions give an indication of the importance and relationship between them in language program evaluation; the “summative-formative”

dimensions, the “product-process” dimensions, and the “quantitative-quantitative” dimensions. Based on their definitions, the following is how these dimensions are closely linked [4]:

a. The “Summative-Product-Quantitative” Link.

Summative evaluation occurs at the end of a program that has been completed to determine whether the program was successful in achieving its objectives. Therefore, it concerns whether the product has met the requirements or not, regardless of what happened during the program. The measurement of the terminal objectives of a program is also product oriented and summative in nature. It is almost always quantitative in nature, such as course achievement results and end-of-course questionnaire responses. Therefore, the “summative-product-quantitative” dimensions can be interrelated along the same form and feature.

b. The “Formative-Process-Qualitative” Link.

A formative evaluation takes place during the program development, implementation and operation. The aim is to find out what is working well, and what is not, and what problems need to be addressed [6, 17]. It focuses on on-going development and improvement of the program. It is concerned with gathering information regarding the processes in a program to gain an understanding of what happens throughout a program and why things turned out the way they did. Information of this kind is usually obtained through qualitative methods such as observations, interviews and document analysis. As such, the formative dimension is closely associated to the features of the process and qualitative dimensions.

The opposing dimensions for each area (summative vs formative, product vs process, and quantitative vs qualitative) have been considered dichotomies, i.e. separate entities that are mutually exclusive [1]. In the practical sense, this would mean that in formulating an evaluation, one has to be selective over the other. However, recent experiences have shown that in addition to understanding the differences and similarities between them, they can be tailored to complement each other to suit a particular evaluation exercise based on the purpose and the type of information required. An evaluation should utilize both points of view, where all available perspectives may prove valuable in terms of enhancing the value of its measurable outcomes, at the same time eliminating bias and strengthening conclusions [1, 4, 6, 18]. The preoccupation to distinguish the different methodologies and approaches are no longer real concerns, but instead should be adapted to capitalize on the strengths of each dimension to maximize the yield of both types of information [1, 12, 20].

However, in deciding on the different types of procedures to be selected in the data collection in this study, the author used Table 1 as a guiding principle. This is because different types of procedures involve different issues depending on program conditions, such as the needs of the stakeholders, access to program information and participants, as well as time limitations.

Table 1 Some Questions Relating the Possible Choice of Evaluation Dimensions to Program Conditions

Issue	Questions to be answered in selecting a dimension(s) for evaluation
The needs of the stakeholders	Why is the evaluation being conducted? What information is being requested and why?
Access to program information	What kind of data can be made available to the evaluator by the program institution? Will the evaluator be granted access to documents pertaining to the program?
Access to program participants	Will the evaluator be granted access to interview program participants? Will the evaluator be granted access to carry out classroom observations? Will the evaluator be available at the disposal of the program for the purpose of evaluation throughout its duration?
Time limitations	When will the evaluator be allowed to begin the evaluation at the program location (point of entry)? Will the program participants be granted time to be interviewed by the evaluator? When will the evaluation report be required? (e.g. - is there enough time to gather, analyze and interpret quantitative/ qualitative data?)

Source: [8, 10, 18, 21, 22, 23]

Finally, after judicious considerations, the author used the data gathering procedures as shown in Table 2 which managed to combine all the dimensions of evaluation in order to maximize the evaluation findings and to address questions not only on the program effectiveness but also to make highly informed recommendations for program improvement.

Table 2 the Multi-Dimensional Data Gathering Procedures

Data Gathering Procedures	Dimensions of the Evaluation		
	Summative/ Formative	Product/ Process	Quantitative /Qualitative
Classroom observations of teachers using a checklist at different stages of the program	Formative	Process	Qualitative
Questionnaires which included both closed-ended and open-ended questions at the end of the program	Summative	Product	Quantitative/ qualitative
Interviews with teachers at the end of the program	Summative	Process/ Product	Qualitative
Interviews with students at the end of the program	Summative	Process/ Product	Qualitative
The evaluation of the teaching material using a checklist with a rating scale	Summative	Product	Qualitative

8. CONCLUSION

The experience of the author showed the benefits of a "tailor-made" multi-dimensional approach in conducting the evaluation study. Firstly, it enabled the author to gain an in-depth understanding of the theoretical foundations of the evaluation process. As a language program is a complex interaction of audience, goals, context, and themes [1, 10], it also allowed the author to be more flexible and holistic in his investigative techniques by combining empirical precision with descriptive precision [24]. In the process, the qualitative data enhanced the value of numerically represented findings from the quantitative method through rich descriptions and by unveiling underlying reasons to explain outcomes and relationships. In addition, the qualitative element allowed the researcher to focus on certain aspects of interest of the program that he or she could not have achieved by using the quantitative method alone [10]. Therefore, the use of both quantitative and qualitative data provided a richer set of information for decision-making and gave a better understanding of the research problems in addressing the related issues of program effectiveness and suggestions for improvement [10, 19]. It can be determined that in order to obtain highly accurate conclusions that lead to credible judgments, evaluations must incorporate multiple procedures for gathering information which involve the different dimensions. This is because language programs are complex and contain a diversity of features which require different types of information from different sources. This enables particular items of interest or phenomena to be explored in different ways and 'triangulation' can be applied, where findings can be cross-checked across methods and sources for enhanced accuracy and validity [4, 6, 8, 10, 19, 25]. It can be established that the inclusive and multi-dimensional data gathering approach is able to provide the relevant stakeholders with highly valid, reliable and credible accounts.

9. REFERENCES

- [12] Alderson, J.C. 1992. Guidelines for the evaluation of language education. In Alderson, J.C. & Beretta, A. (eds.). Evaluating second language education, pp. 274-304. Cambridge: Cambridge University Press.
- [21] Alderson, J.C. & Scott, M. 1992. Insiders, outsiders and participatory evaluation. In Alderson, J.C. & Beretta, A. (eds.). Evaluating second language education, pp. 25-57. Cambridge: Cambridge University Press.
- [18] Beretta, A. 1992. Evaluation of language education: an overview. In Alderson, J.C. & Beretta, A. (eds.). Evaluating second language education, pp. 5-24. Cambridge: Cambridge University Press.
- [1] Brown, J.D. 1989. Language program evaluation: a synthesis of existing possibilities. In Johnson, R.K. (Ed.). The second language curriculum (222-241). Cambridge: Cambridge University Press.
- [19] Creswell, J.W. 2005. Educational research. New Jersey: Pearson.
- [16] Graves, K. 2008. The language curriculum: A social contextual perspective. Language Teaching 41(2): 147-181.

- [25] Genesee, F. & Upshur, J.A. 1996. Classroom-based evaluation in second language education. Cambridge: Cambridge University Press.
- [11] Hedge, T. 2000. Teaching and learning in the language classroom. Oxford: Oxford University Press.
- [17] Jarvis, G.A. & Adams, S.J. 1979. Evaluating a second language program. Virginia: Center for Applied Linguistics.
- [13] Johnson, R.K. 1989. A decision making framework for the coherent language curriculum. In Johnson, R.K. (ed.). The second language curriculum, pp. 1-23. Cambridge: Cambridge University Press.
- [9] Kiely, R. 2013. The Role of Programme Evaluation in Curriculum Development OR 'How Are We Doing?'. In Powell-Davies, P. & Gunashekar, P. (eds.). English language Teacher Education in a Diverse Environment, pp. 122-128. British Council.
- [4] Kiely, R. & Rea-Dickins, P. (2005). Program evaluation in language education. Hampshire: Palgrave Macmillan.
- [14] Low, G. 1989. Appropriate Design: The internal organisation of course units. In Johnson, R.K. (ed.). The second language curriculum, pp. 136-154. Cambridge: Cambridge University Press.
- [10] Lynch, B.K. 2003. Language assessment and program evaluation. Edinburgh: Edinburgh University Press.
- [22] Lynch, B.K. 1996. Language program evaluation. Cambridge: Cambridge University Press.
- [23] Mitchell, R. 1992. The "independent" evaluation of bilingual primary education: a narrative account. In Alderson, C. J. & Beretta, A. (eds.). Evaluating second language education, pp. 100-136. Cambridge: Cambridge University Press.
- [2] Muhammad Salim Tufail. (2010). An Evaluation of the Malaysian Armed Forces General English Upper Intermediate Course (Officers). (PhD thesis). Universiti Kebangsaan Malaysia.
- [20] Murphy, D.F. 1985. Evaluation in language teaching: assessment, accountability and awareness. In Alderson, J.C. (ed.). Evaluation. pp. 1-18. Oxford: Pergamon.
- [5] Norris, J.M. (2016). Language Program Evaluation. The Modern Language Journal, 100 (Supplement 2016), 168-189.
- [7] Nunan, D. 2003. Methodology. In Nunan, D (Ed.). Practical English language teaching (3-22). Boston: Mc Graw Hill.
- [15] Nunan, D. 1989. Understanding language classrooms. New York: Prentice-Hall.
- [24] Onwuegbuzie, A. J. & Leech, N. L. 2004. Enhancing the Interpretation of "Significant" Findings: The Role of Mixed Methods Research. The Qualitative Report 9(4): 770-792.
- [3] Richards, J.C. 2002. Program factors in effective foreign and second language teaching (online) <http://www.professorjackrichards.com/pdfs/program-factors-effective-teaching.pdf> (12 February 2007).
- [6] Richards, J.C. (2001). Curriculum development in language teaching. Cambridge: Cambridge University Press.
- [8] Weir, C. & Roberts, J. (1994). Evaluation in ELT. Oxford: Blackwell.