An Introduction to Course and/or Program Evaluation

Mohammad Zohrabi  
*The University of Tabriz*


Program and/or course evaluation is, in fact, one of the essential aspects of any curriculum. It is a kind of quality control in which various aspects of an instructional program are explored. Program evaluation is an attempt in which different elements of a given curriculum are scrutinized in depth. To this end, an evaluator makes every effort to collect information from different sources such as students, teachers, administrators, course designers, program staff, and so on. Also, the evaluator tries to gather data through different procedures such as questionnaires, interviews, observations, course documentations etc. Collecting hybrid type of information helps increase validity and reliability of the study. The main concern of program evaluation is to ensure that acquisition is taking place, teaching techniques and strategies are useful, materials are relevant and interesting, resources are available and adequate, and so forth. In the beginning, program evaluation was merely quantitative and/or experimental in nature. However, with the passage of time it was realized that qualitative and/or natural approaches are more useful than laboratory-like artificial experimental methods. The important issue in any course evaluation is to bring about some necessary changes. This article tries to elaborate on the preliminary aspects of program evaluation.

**Key Words:** evaluation, testing, objectives

1 Introduction

This article comprises five sections in which the following topics will be elaborated on. At first, an overview and development of program evaluation within the context of language teaching will be discussed. Then, some definitions of program evaluation will be presented. Also, the role and purpose of evaluation and its importance will be touched on. One of the important branches within the evaluation process is the testing issue. Therefore, it will be tried to allude to the differences between evaluation and testing. Finally, evaluation is an issue which different researchers have different views about it. Therefore, the opinions concerning its acceptance and rejection will be commented on.
2 What Is Evaluation?

Evaluation is a key element in any educational endeavor especially within the curriculum development (Worthen et al. 2003). Most of us either as language teachers or students are familiar with program evaluation and have been involved in it in one way or another. However, evaluation has received scant attention in the field of foreign or second language teaching. Unfortunately, in the field of language teaching very few books or journals have appeared on evaluation. But in the fields of social and educational sciences many scholarly published materials proliferate.

It goes without saying that the most influential scientist in the field up till now has been Ralph Tyler (1949). In Tyler’s approach the predetermined issues are compared with the existing issues. At first, behavioral goals are identified, and then tests are developed based on those goals. This approach has some defects. For instance, there are some unpredictable issues that are abstract and can hardly be determined and defined. Therefore, limiting evaluation to only behavioral objectives leaves out these issues. Also, one of the main deficiencies of Tyler’s approach is its lack of attention to process. It practically disregards what actually takes place within an instructional program.

Evidently, interest in the evaluation process mainly came into prominence in the 1960s. Language teacher employs different evaluation techniques in order to assess the students’ progress or the course’s success (Fraenkel & Wallen 2003). In the early days evaluation was thought of as the testing of students through the end-of-semester tests. The emphasis was on the final product, that is, the students’ test results. There was no attention being paid to the process of evaluation. In this regard Lewkowicz and Moon (1985, p. 46) criticize these kinds of measurements and state that “They [teachers] have not generally been concerned with whether the learners were able to interpret the results in a way which would enable them to modify their behavior; nor have they taken into account their feelings or reaction to the evaluation procedure themselves.”

It can be argued that program evaluation consisted of and was equal to investigating the efficiency and working of language teaching methods and materials. Some of these studies which concentrated on comparing and contrasting different methods and approaches of classroom language teaching are as follow. The first study is about Raymond F. Keating’s (1963) experiment on the effectiveness of laboratory use. Keating tested both the experimental students who were taught through laboratory and the control group who received regular treatment by reading, listening and speaking tests. However, the control group students outperformed the laboratory or the experimental students. Nonetheless, this study was criticized for its lack of validity, lack of controlling various variables and faults in its experimental design.
Another study concerns the usefulness of teaching methodologies. It was a two-year study carried out by Scherer and Wertheimer (1962) which became known as the Colorado Project. In this study two groups of students were taught German language through audio-lingual and grammar-translation methods. Then, the students were tested at the end of the course through achievement tests. The experimental group who was taught through audio-lingual method performed well in listening and speaking achievement tests. However, the students who were taught through grammar translation method did better in translation, reading and writing. This study revealed that the students’ achievement depended on the methods that they were being taught. This investigation also received several criticisms with regard to its lack of experimental control, its small size and lack of description of the teaching procedures. The shortcomings of this quantitative and/or experimental study motivated later studies to consider the process, i.e. the qualitative methods rather than the product of teaching programs.

Chastain and Woerdehoff (1968) conducted an entirely controlled experimental method study. They too investigated the efficacy of audio-lingual and traditional methods. They referred to the traditional approach in their study as the cognitive code method. In order to control different variables, they used pretest and a posttest was administered at the end of the course. The researchers concluded that the cognitive code method was more superior to the audio-lingual method. Generally, one evaluation model which was in vogue in the 1970s was Stufflebeam et al. ’s (1971) CIPP evaluation (Context, Input, Process and Product). In this model context evaluation is concerned with the evaluation of a program of study to explore its strengths and weaknesses in order to make it better. Input evaluation determines the available means through which to evaluate the objectives of a program. In input evaluation the evaluator attempts to use every resource to carry out the evaluation. Process evaluation is the actual phase of fulfilling an evaluation. The evaluator employs several techniques in order to execute the evaluation and in this way works toward the improvement of the program. In the end, the product evaluation tries to assess how far the program objectives are attained. On the whole, this model attempts to supply information for decision-makers.

All of the above-mentioned studies indicate that in the early days the investigations were rather concentrated on the use and differences of various methods of language teaching. There was hardly any attention being paid to the process of language learning. Most of those studies were using quantitative methods and data for interpreting their results (Zohrabi, 2008). Little by little, however, method and experimental studies gave way to more naturalistic studies of classroom processes. Royse et al. (2005) believes that the interest in more naturalistic data gathering and interpretation began in the 1980s. For instance, Guthrie (1982) studied a language teaching program in California in which she used an ethnographic approach. She observed the
actual process of language teaching which took place in the classroom. She also interviewed several students, the teacher and administrators and studied program documentations as well as field notes. This investigation was truly a qualitative study in which the students’ and the teacher’s attitudes toward the course, the methodology and materials were scrutinized.

Generally, the attitudes toward the program evaluation have changed in recent decades and the evaluators have many different research designs at their disposal. They might want to concentrate on negotiation, policy and a range of other issues. More importantly, they can gain more utilisable information in a short period of time (Posavac & Carey, 2002). By and large, course or program evaluation can take place at different levels. Langbein and Felbinger (2006) surmise that it may occur at two levels: macro-level and micro-level. At macro-level a program is usually evaluated by experienced researchers along with several personnel who are expert in the field. This type of evaluation involves a large-scale national and/or state program involving students, teachers, administrators and so on. In these studies the whole process and details of a program are investigated. On the other hand, a micro-level or small-scale evaluation might involve only the teachers and learners. Usually, in this type of study only one classroom is investigated by its own teacher.

3 Approaches to Evaluation

Different experts have rendered varied definitions on evaluation. A rather more elaborate and comprehensive definition is presented by Brown (1995). Brown’s focus of attention is on a specific type of curriculum that is being evaluated. Thus, he (ibid. 24) affirms that “Evaluation might be defined as the systematic collection and analysis of all relevant information necessary to promote the importance of the curriculum and to assess its effectiveness within the context of the particular institutions involved.” Nevertheless, Lynch (1996, p. 2) suggests a rather condensed and straightforward definition of evaluation and puts it as follows: “Evaluation is defined here as the systematic attempt to gather information in order to make judgments or decisions. As such, evaluative information can be both qualitative and quantitative in form, and can be gathered through different methods such as observation.”

Ryan (2007) holds that evaluation is concerned with process of description and making value judgments. He reckons that evaluation is an indispensable part of a syllabus and it should be included in any curriculum. In this way evaluation can be made accountable to the changing needs of a society. Robinson (2003, p. 199) presents a succinct definition of evaluation as “the collection, analysis, and interpretation of information … for forming judgments about the value of a particular program.” Therefore, Murphy (1985, p. 4) argues that “evaluation should be an integral part of the working
An Introduction to Course and/or Program Evaluation

of the curriculum to ensure that what is done is worthwhile, necessary and sufficient.” On the whole, emphasis should be on evaluating the entire learning process rather than the learners. It is believed that the emphasis should be on investigating the productivity of a course rather than merely on assessing the learners. Assessing learners barely provides us with necessary data. Therefore, stress should be on the process of course evaluation. Hutchinson and Waters (1995, p. 144) hold that “This kind of evaluation helps to assess whether the course objectives are being met – whether the course, in other words, is doing what it was designed to do.” There are yet other researchers who assume that the purpose of evaluation is to inquire into the usefulness and quality of a program. For instance, Jordan (1997, p. 85) contends that: “The purpose of evaluation is to collect information systematically in order to indicate the worth or merit of a program or project … and to inform decision making both from the point of view of development and the end product.”

More importantly, Dudley-Evans and St John (2000) maintain that evaluation is concerned with collecting necessary data and making decisions based on that data. They (ibid., p. 128) state that “For us, evaluation is a whole process which begins with determining what information to gather and ends with bringing about change in current activities or influencing future ones.” However, evaluation is not merely the process of gathering information. It is studying the shortcomings of a course of study in order to see whether the objectives have been achieved, so that to make appropriate decisions according to that information. Nunan (1999b, p. 185) presumes that “The data resulting from evaluation assist us in deciding whether a course needs to be modified or altered in anyway so that objectives may be achieved more effectively.” It should be noted that evaluation is not a one-way process in which only the learners are judged by outsider researchers. It is, however, a multi-way process in which the learners, the learning process, the teachers, and the whole teaching-learning process are scrutinized. To this end, Lewkowicz and Moon (1985, p. 47) assert that “Evaluation, therefore, in our opinion should be viewed both as an ability to self-assess and critically judge one’s own learning and performance and an ability to understand, learn from and utilize feedback and evaluation from a variety of other sources.”

In sum, evaluation is a process in which a great amount of data is gathered through various procedures from different sources. Then, this information is interpreted to make some important decisions based on the research results. These decisions might require a change and effect a drastic alteration in the outline and process of a language program instruction. All these efforts are made to the betterment of a course of study and bringing about satisfactory results (Zohrabi, 2009). Therefore, “Program evaluation which describes the achievements of a given program, provides explanations for these, and sets out ways in which further development might be realized” (Kiely, 2009, p. 99).
4 Objective of Evaluation

Any language teaching program comprises several courses. These courses invariably have some common goals. Therefore, any educational program should include evaluation as an integral part in its design. As a matter of fact, evaluation should be considered as a vital element in any learning and teaching situation. It should not be regarded as an extra feature imposed on the program. Fitzpatrick and Edwards (2002) is convinced that evaluation is an essential part in the instructional process and we should be careful to ensure that it produces the intended outcomes. As Parilah and Zohrabi (2008) put it the first priority in any evaluation endeavor is to elucidate the objectives. That is, it should be made clear why the evaluation is being carried out. For the sake of program evaluation Langbein and Felbinger (2006) surmise that the goal is to measure the efficiency and to compare its quality with other identical language instructional programs. The important point for Flowerdew and Peacock (2001) is the result of the evaluation which should bring about some necessary recommendations for change.

The vital issue in any educational program is the degree to which the students learn the materials. Therefore, Dudley-Evans and St John (2000) are of the opinion that evaluation should take care of students’ acquisition and assess whether the resources have properly been utilized. Langbein and Felbinger (2006) hold that information that is gathered through evaluation process may be used to make any requisite revisions of a course, or it may be employed to apply to other similar instructional courses. These authors also express that the evaluation outcomes can be utilized in producing course specific materials and in this way a great deal of time can be saved. McDavid and Hawthorn (2005) argue that program evaluation can identify the reasons for the failure of the courses and in this way provide some necessary solutions.

It cannot be ensured that evaluation would unveil every defective point and assist the evaluator in determining every minute detail. It is almost irrational to think too much of evaluation and invest too much on it. Alderson and Beretta (1996, p. 273) maintain that evaluation “is not an exact science.” On the whole, it can be stated that there are so many variables involved in the evaluation process so that its success quite fairly depends on the amount of cooperation that it receives. Beretta (1996, p. 265) notes that “Evaluations take place in the real world and not in laboratories, and thus the results are always tentative, judgmental and incomplete.” Therefore, the evaluator should be careful to control and take heed of every variable that is involved as far as possible. The optimal results can be achieved by using different methods of gathering information and, of course, from different sources. The more data and information the evaluator can gather, the more reliable and valid the results can turn out.
5 Terminological Differences

It is clear that there are major disparities between evaluation and testing in language teaching. Evaluation is a cover term which either may or may not include testing (Lynch, 2003). Regrettably, some teachers assume that evaluation means testing the students at the end of the language course and use these terms interchangeably. Murphy (1985, p.10) stresses this issue and asserts that “However, procedures and methods other than for testing learners are not described; the necessity for evaluation is not understood and recognized.” He, then, emphasizes the importance of carrying out evaluation in order to investigate language teaching courses and textbooks.

In testing only the learners’ attainment of course materials, objectives and their mastery of language are assessed (Bachman & Palmer, 2000). At this point it is preferable to notify that testing pertains to the processes that are entirely dependent on assessing the students through various types of tests. Testing, in fact, is a one-way procedure in which the focus of attention is merely on the learners, and the information gathered is quantitative. Additionally, there is no attempt to gather qualitative data concerning the learners’ and the teachers’ perceptions concerning materials, methodology, needs, objectives, and so on. However, evaluation is an extensive area in which various types of information may be gathered (McDavid & Hawthorn, 2005). Within evaluation practice, though, the evaluator may prefer to collect some information through testing process. Then, it is safe to assume that testing is one of the branches within the evaluation process. At this juncture, Hudson (1989, pp. 262-263) demonstrates that pencil-and-paper testing is “non-authentic” and takes language away from its real context. Therefore, he emphasizes that “the requirement for authentic evaluation would disallow testing setting. Other forms of measurement would be needed” (ibid.).

In a learning-centered approach to language instruction, the contributions of the learners have been emphasized. Therefore, exploring learners’ views is an essential asset for the evaluator. It can be reasoned that an evaluator could opt for various methods and means of collecting data in which testing may rather play a very minor role. Lewkowicz and Moon (1985, p. 47) acknowledge that “It [evaluation] no longer narrowly applies to testing but encompasses a much wider range of concerns.” Moreover, Avery et al. (2004) affirm that in a learner-centered approach to learning, evaluation is an ongoing process which is carried out by its members, i.e. learners and teachers. Alderson (1996, p. 284) severely criticizes the use of tests in evaluation and emphasizes that “Evaluation can be carried out without tests, and no doubt often should be. Indeed, in recent years there has been a tendency to question the automatic inclusion of language tests in the evaluation of language education. Testing has many opponents and critics,
and it is common to argue that since tests cannot tell one everything one might wish to know about learning, thus should not be used at all.”

To be on the safe side, Creswell (2003) emphasizes that the evaluators had better gather different type of information from different sources and, of course, use different procedures of data collection. It can be inferred from the above discussions that the evaluators would rather not rely too much on testing the learners in the program evaluation. It is because testing results may not provide the necessary and adequate information for the evaluation purposes.

6 The Role of Evaluation

Obviously, different teachers have naturally different attitudes toward evaluation. If an evaluator tended to evaluate one of our language teaching classes what our reaction would have been? Alderson and Beretta (1996, p. 273) contend that it “can be face-threatening, and we need … to reduce its threatening nature, and make it more acceptable to all concerned.” Mainly, some teachers do not want to be judged by outsider evaluators or other teachers. Therefore, evaluators should take heed of the specific site where they carry out their evaluation. In this regard, Hudson (1989, p. 259) affirms that “the context of evaluation should always be kept in mind because evaluation always has some consequences.”

Some teachers, by and large, do not like to be observed and “too often it is assumed that any resultant evaluation document will be negatively critical” (Robinson, 1991, p. 69). Also, some curriculum developers feel that their interests are threatened by the evaluation and because of this “evaluation tends to be neglected in project planning and development” (Alderson & Scott, 1996, p. 27). However, sometimes evaluation is inescapable and is implemented “largely because of outside pressure on educators to explain and justify what they are doing” (Murphy, 1985, p. 2). In this case, when the evaluation is imposed on a course by some outside political pressure it “is then perhaps necessarily perfunctory, superficial and inadequate” (Alderson & Scott, 1996, p. 27).

An evaluator ought to ensure the teachers that evaluation “is one of the most useful means of obtaining information about what is working and what is not” (Nunan, 1999, p. 147). The evaluator should make every effort to gain the confidence of students, teachers and administrators. It is believed that “Confidence will increase if the evaluation is seen to have beneficial results, for example, an improvement in teaching conditions, rewards for endeavors etc.” (Robinson, 1991, p. 69). The important point in every educational program is that the teachers are better to be encouraged to carry out a self-analysis of their own classroom work. This self-evaluation will, in consequence, result in self-development and in turn will contribute to curriculum development (Bailey, 2006).
An Introduction to Course and/or Program Evaluation

Those being evaluated might half cooperate and might not reveal their real views. Performing an evaluation may disturb and interfere with the normal or regular schedule of a teaching program. Therefore, the evaluator might not receive complete cooperation of the whole parties involved in it. These partial co-operations stem from the lack of time, interest etc. (Mason, 2007). However, there is rarely any need to be disappointed because “evaluations receive cooperation from all parties and provide useful information to insiders on how their work can be improved, while offering accountability to outside stakeholders …” (Long & Richards, 1996, p. ix). It is clear that generalizations are difficult to make and the amount of cooperation which the evaluator might receive depends on the single individuals who are being evaluated and the particular institution’s rules and regularities.

7 Conclusion

It is safe to say that program and/or course evaluation is a kind of glue which joins all the elements of a curriculum together. Without program evaluation we cannot make sure whether the students true needs are met, whether they are satisfied with the course under study, whether the materials are effective and whether testing motivates more learning or hinders it. Therefore, program evaluation attempts to investigate a course from different perspectives. The important point is that an evaluator cannot properly succeed to carry out an evaluation unless different parties involved in it try to help him/her in one way or another. Program evaluation, in fact, is a vast and broad endeavor. It requires a great amount of time, energy, expertise, experienced personnel, resources and so on. It cannot be performed at a given point in time. It is an ongoing process which begins at the start of the course, continues till its end and even after it. Program and/or course evaluation tries to bring about some necessary and adequate modifications. Therefore, the end result or outcome of an effective evaluation process is to motivate some useful improvements in a course of study.

References

Mohammad Zohrabi


An Introduction to Course and/or Program Evaluation


Mohammad Zohrabi


Mohammad Zohrabi
Department of English Language,
Faculty of Literature,
The University of Tabriz, Tabriz, Iran
Phone: 009841133921692
E-mail: mohammadzohrabi@gmail.com

Received: July 24, 2011
Revised: November 15, 2011
Accepted: December 7, 2011