

Journal of English Language  
Teaching and Learning  
University of Tabriz  
No. 20, 2017

## Goals, Dimensions, and Design of Program Investigation\*

**Leila Dobakhti\*\***

Assistant Professor of TEFL, Tabriz Islamic Art University  
(corresponding author)

**Mohammad Zohrabi**

Assistant Professor of TEFL, University of Tabriz

### **Abstract**

If we are keen to boost the process of language learning, we need to study every aspect and component of our course. To this end, we carry out an investigation in which every detail of the course is put under microscope. Assessment of a course is an attempt in which different type of information is gathered systematically in order to study the working of a language instruction program. Certainly, program evaluation might be carried out for fulfilling different goals and purposes. Therefore, an evaluator should clarify the audience and the objectives of conducting program evaluation from the very beginning in order to obtain some adequate results. Different researchers have presented different proposals for the specific time of collecting data, type of data, and procedures of gathering data. Consequently, an evaluator should be familiar with the different dimensions of program evaluation. Also, some researchers treat program evaluation as a neat product, i. e. merely consisting of learners' end of instruction scores. Yet, others consider program evaluation as a process. At this juncture, the important issue is the selection of an appropriate design for the implementation of the actual process of the evaluation. It is suggested that the *illuminative model* is rather a proper design for program evaluation because of its different advantages. This article attempts to shed some light on the dimensions, goals, and design of the course study.

**Keywords:** course study, dimensions, illuminative model

---

\* Received date: 2017/04/07      Accepted date: 2017/11/10

\*\* E-mail: l.dobakhti@tabriziau.ac.ir

### **Introduction**

To operationalize the process of course evaluation, three important features of a program will be elaborated on. These characteristics include audience and their goals, dimensions, and design. Mainly, the audience and the objectives of the evaluation should be determined prior to the evaluation process. The purpose of the study is very crucial because it determines the dimensions and the design of the study. Also, before embarking on the actual process of conducting the evaluation endeavor, the evaluator should determine the dimensions of his/her research. These dimensions shape the overall frame of the study. On the whole, the design of the study plays a crucial role in the evaluation process. It is the design of the research which determines the type of data collection and analysis.

### **Goals and Audience**

The first stage of an evaluation is very important because it is at this stage which the audiences and their goals of conducting evaluation are determined. In every evaluation there are different audiences with different purposes. To clarify who the audiences are, Lynch (1996, p. 3) raises two key questions: “Who is requesting the evaluation? Who will be affected by the evaluation?” The first question might be put forward by the funding agencies who want to see value for money. These stakeholders want to know whether or not the students’ test scores are high enough and satisfactory in order to continue supporting the program financially. The second stakeholder might be the institution itself which may want to see whether or not the program is successful. The second question is related to the students and to some part to the teachers who might be affected by the results of the evaluation. Therefore, the audiences of an evaluation might include funding agencies, educational institutions, administrators, program staff, curriculum developers, parents, teachers, other researchers and evaluators, and students who might be interested in the evaluation results.

In order to determine the goal(s) of an evaluation, Lynch (1996) sets forth two questions: “Why is the evaluation being conducted? What information is being requested and why?” It is clear that different stakeholders have different goals. After the audiences of the evaluation

have been determined, the evaluator attempts to reach an agreement among them in order to conduct the evaluation. At this point there are various points of view and, of course, expectations. Obviously, the evaluator cannot carry out the evaluation process according to his/her own objectives. As Alderson (1996) points out, every stakeholder has his/her own goals and even these goals are sometimes in conflict with each other. Mainly, the evaluator should clarify the specific audiences and their particular purposes. Therefore, before conducting an evaluation, the evaluator should try to find out answers to these questions: “Why am I doing this research? In whose interests is this research?” (Blaxter *et al.*, 2006, p. 10-13). However, Brown (1989, 1995) considers goals of fulfilling an evaluation from four different points of view. The following section explicates fairly in more elaborate detail Brown’s goals of approaching program evaluation.

An evaluation can be approached from different perspectives based on a program’s goals and audiences. Generally, there are several factors which influence a choice of one approach over another. Mainly, the key elements which affect our choice include the audience, objectives and goals of the research problem, to name but a few. Therefore, according to a program’s particular goals, Brown (1989, p. 224 & 1995, p. 219-20) offers four approaches to program evaluation. Therefore, any program evaluation might be carried out based on one of the following approaches:

- 1- Goal-attainment and/or product-oriented approaches
- 2- Static-characteristic approaches
- 3- Process-oriented approaches
- 4- Decision-facilitation approaches.

#### **Goal-Attainment and/or Product-Oriented Approaches**

The main focus of these approaches is on the goals and objectives of a language instruction program. These approaches intend to investigate whether the overall objectives are or are not achieved. One of the chief advocates of this approach was Tyler (1942). He believed that any program should have specific aims and quantifiable behavioral topics. Therefore, the goals of the program should be measured at the end of the instruction in order to see whether they are or are not attained.

In Tyler's approach there is no room for qualitative and/or non-quantifiable information. There is also no consideration being paid to the actual classroom process of language learning and teaching. The main factors in this approach are the measurable products and behaviors. In fact, this approach only assesses the students' attainment of course objectives. Also, the information gathered at the end of the program cannot be fed into the course and has rarely any feedback usefulness. Moreover, there is hardly any data collected on the perceptions' of students, teachers, and administrators. Therefore, product-oriented approaches have relatively many shortcomings which little by little have lost their credibility in the last decades. It can be stated that this approach can hardly be considered as an appropriate way of implementing the program evaluation. This approach is merely testing the students' attainment of course objectives which is usually fulfilled by any language teacher.

#### **Static-Characteristic Approaches**

The second type of approach – static-characteristic approaches – proposed by Brown (1989, 1995) relates to the resources provided by an educational institution to its students and instructors. In this approach an outsider evaluator usually visits the site in order to make a report on the different static characteristics of the physical plant. That is, the evaluator observes and records the number of library books, lab tapes, CDs, learning aids, computers, number of classrooms, number of masters and PhDs among the staff, the ratio of learners to instructors, availability of self-access learning center, number and varieties of scholarly journals, photocopy machines, etc. Therefore, if the so-called institution already holds these resources, it can receive an accreditation.

This type of evaluation has nothing to do with the process of learning and teaching. It deals, in fact, with non-learning and teaching aspects of an institution rather than what is happening inside its classrooms. Although this type of evaluation is always crucial to ensure that the students have access to various necessary resources, it should be complemented with more comprehensive evaluations carried out on the teaching and learning processes. To a large extent, the necessity of these resources cannot and should not be ignored. Every educational institution must provide its students and instructors with enough

journals, books, CDs, tapes, self-access learning center, copy machines, computers, and so on. Obviously, these equipments and means of learning affect the amount and rate of learning substantially. The lack of these resources not only frustrates the teachers but also the students.

### **Process-Oriented Approaches**

The third type of evaluation approaches proposed by Brown (1989, p. 226) is the process-oriented approaches. In this type of evaluation the actual processes of learning and teaching are evaluated. Unlike the static-characteristic approaches which investigate non-teaching-learning, the process-oriented approaches explore what is happening not only inside but also outside the classroom. Here the evaluator makes every effort to examine the components of the curriculum such as materials, testing, needs analysis, teaching activities, and so forth. It is realized that in order to make any necessary modifications and changes in the curriculum, the whole program should be scrutinized. Process-oriented experts argue that program evaluation is an ongoing process which cannot be performed at a specific point in time. The information should be gathered from the start of the course in order to provide feedback to its betterment and also to be used for the subsequent similar programs. To this end, observing classroom activities and interviewing students, teachers and administrators can give invaluable information to the evaluator. One of the chief proponents of this approach is Scriven (1967). To this end, Brown (1989) mentions some of Scriven's contributions to this approach. Therefore, drawing on Brown some of Scriven's accomplishments can be summarized as follows:

- the necessity of evaluating the objectives and studying their worth at the same time.
- the importance of investigating the expected objectives and at the same time to be prepared to take into account unexpected results.

### **Decision-Facilitation Approaches**

The last approach suggested by Brown (1989, p. 227 & 1995, p. 219-20) is concerned with the decision-facilitation approaches. The main purpose of this approach is to gather data and then pass on this information to the decision makers to make any necessary decisions. In

this approach the evaluator tries to collect as much information as possible and at the same time attempts to avoid any specific interpretation of them. The main purpose of this approach is to facilitate the process of decision-making. Therefore, it is the program administrators who interpret the data and make any necessary decisions and judgments.

However, sometimes it is difficult for the evaluator to merely gather data and not to make any interpretations of them. More importantly, the evaluator who collects the information is rather in a privileged position to make practical and reasonable interpretations of the data. Therefore, it can be implied that not involving the evaluator in decision-making process decrease the credibility of the decisions reached by the administrators.

### **Dimensions of Program Evaluation**

When planning to fulfill an evaluation and after determining the particular approach, the evaluator should consider three dimensions which could help shape the perspectives of the evaluation. There are several dimensions which might help the evaluator to focus on the specific type of information that is to be gathered. These dimensions are one of the important characteristics of any evaluation and form an outline for the evaluator to implement the evaluation activity. To this end, Brown (1989, p. 229 & 1995, p. 228) presents three important dimensions of evaluation and contends that they are complementary rather than dichotomous.

- Formative vs. Summative
- Process vs. Product
- Qualitative vs. Quantitative.

### **Formative vs. Summative Evaluation**

The distinction between *formative* and *summative* evaluation was made for the first time by Scriven (1967). On the whole, formative evaluation is ongoing and takes place during the life of a course of study and in this way it is used in “improving the program during the course of its delivery” (Nunan, 1999, p. 192). Robinson (1991, p. 65) also expresses that “the results obtained can be used to modify what is being done.” Furthermore, Jordan (1997, p. 89) states that information which comes

about as a result of formative evaluation is related to “the development process” and consequently some “aspects of the course can be adapted in an ongoing way.” Moreover, Dudley-Evans and St John (2000, p. 128) emphasize the importance of formative evaluation in ESP and assert that it “is typically undertaken at intervals and will consist of a series of mini-evaluations.” Finally, Richards and Schmidt (2002, p. 209) put forward a clear and succinct definition of formative evaluation: “the process of providing information to curriculum developers during the development of a curriculum or program, in order to improve it.”

On the other hand, summative evaluation is the process of gathering information at the end of a program. Nunan (1999, p. 192) believes that this type of evaluation is “not capable of improving that particular program” and, therefore, it provides feedback for “the modification or curtailment of succeeding programs.” In this regard, Robinson (1991, p. 65) also accepts that “the course or project is finished and ... it is clearly too late to do any fine-tuning.” Additionally, Dudley-Evans and St John (2000, p. 128) affirm that summative evaluation can even be carried out after the end of a program in order to “determine its longer-term impact.” They also emphasize that summative evaluation provides useful information “for durable courses.” Therefore, the information that results from summative evaluation is used to judge “the degree to which the program was successful, efficient and effective” (Brown, 1995, p. 225). To this end, Jordan (1997, p. 89) proposes that in order to perform summative evaluation “Students will normally be given a questionnaire to complete anonymously.”

The decision about choosing either the summative or formative evaluation “hinges on the purposes for information gathering and on the types of decisions that will ultimately evolve from each purpose” (Brown, 1989, p. 229). However, Robinson (1991, p. 65) is of the opinion that for any program “both types of evaluation can be undertaken.” Therefore, in order to complement each other, it will be better that every curriculum to draw on both types of information gathering activities. However, there are sometimes a little overlapping and it is fairly difficult to draw the lines between them. To this end, Alderson and Scott (1996, p. 39) point out that “no evaluation is purely summative, purely formative.”

### **Process vs. Product Evaluation**

In general, process evaluation concentrates on the actual functioning of a program. Therefore, it focuses on what is happening in the program (Brown, 1989). In this regard, Robinson (1991, p. 65) adds that process evaluation is related to the “teaching and learning strategies or processes, and administrative and decision-making processes.” Furthermore, Richards and Schmidt (2002, p. 422) explain that process evaluation “attempts to measure ... teacher behavior or processes (what the teacher does in class ...).” Mainly, process evaluation tries to study all the aspects of the curriculum, for example, teachers’ methods of teaching, materials, needs, objectives, testing, and perceptions’ of the students, teachers and administrators, etc. Therefore, it can be surmised that process evaluation is a large-scale study of the curriculum, students, teachers, materials, resources, etc. To this end, Murphy (1985, p. 4) holds that “every facet of the curriculum and its functioning may be evaluated, not just the performance of those whose studies are guided by it.” In general, the information that process evaluation provides is invaluable because it is rather comprehensive and includes all the elements and aspects of the program.

On the other hand, product evaluation focuses on whether the objectives of the program have been attained. That is, in this type of evaluation, information is collected in order to check whether the aims of the program are being achieved (Brown, 1995). Therefore, product evaluation tries to investigate “student product such as examination results ...” (Robinson, 1991, p. 65). In product evaluation, unlike the process evaluation, the focus of attention is mainly on students’ outcomes and achievements. To this end, Richards and Schmidt (2002, p. 42) point out that product evaluation intendeds to assess “pupil learning or products (e.g., as demonstrated on a performance on a test).” Normally, the type of data that is being gathered depends on the purpose of the evaluation. If the emphasis of the research is on the process of the program’s working, then process evaluation can be adopted. However, if the focus of the program is on the outcomes, then students’ achievements can be measured.

There are mainly many reasons that a research might avoid using tests in order to study students’ achievements. Generally, the test scores

rather barely indicate anything about the true achievement of the students. Also, some students attend private language institutes at the same time that they take general English course at the university. So that, their achievement mostly comes about as a result of private language institute's impact. To this end, Sharp (1990, p. 132) rightly states that evaluating students through tests has its own drawbacks and affirms that:

- 1- No explanation is supplied for any disparity in results.
- 2- Little help is offered in deciding which parts of a course have been working successfully and which have not.
- 3- Some of the objectives of a course cannot be easily measured by testing.
- 4- The percentage mark achieved by a student after an exam provides only limited information about whether his needs have been met.

#### **Qualitative vs. Quantitative Data**

As a general rule, in order to gather information, the evaluator can opt either for qualitative or quantitative data. As usual, it depends on the purpose for which the evaluation takes place. However, for any evaluation to be rich and have convincing evidence, it is preferable that both types of data to be collected and used.

Generally, qualitative data can be described as the type of information which can be obtained through observations, interviews, open-ended questionnaires, and so on. Alderson (1996, p. 282) believes that exploring "attitudes and opinions are important to the evaluation." He then emphasizes that "observation, recording and interpretation of events, activities and feeling of participants" (ibid. 283) are also of paramount importance. In addition, Brown (1995, p. 227) assumes that qualitative data contains "more holistic information ... that may not readily lend themselves to conversion into quantities or numbers." Nevertheless, Richards and Schmidt (2002, p. 435) assert that qualitative data "can often be converted into quantitative form." On the whole, in order to interpret the qualitative data, the evaluator can utilize them "in a principled and systematic manner [because] they are more important to actual decisions made in a program" (Brown, 1989, p. 232).

In contrast, quantitative data are countable pieces of information which are usually numerical in form. They can be obtained through “tests and objective-questionnaire” (Dudley-Evans & St John, 2000, p. 128).

Qualitative data are mostly obtained through observations, interviews and open-ended questionnaires and quantitative data are gained through closed-ended questionnaires. A researcher might not use testing procedures to collect data because as Elley (1989, p. 270) affirms “Samples are biased or unmatched ... contamination occurs between experimental and control groups, and tests prove too difficult or too easy for students.” Also, it can be stated that though assessment may be important, “they are not the exclusive, or even the primary, focus of all evaluations” (Nunan, 1999, p. 190). Moreover, the information that can be gathered through tests has rather limited interpretational value because “They provide answers to what questions but cannot easily address the how or why” (Dudley-Evans & St John, 2000, p. 128). More importantly tests assess what teachers think are the objectives of the course, however, many vital and unexpected issues surface incidentally in the actual teaching-learning processes. To this end, Slimani (1996, p. 199) reasons that: “However, since we are concerned with relating learning outcomes to their immediate and potentially determining environment, it appears rather difficult to think of ways of getting at learning evidence through testing and elicitation procedures as traditionally understood.”

### **Research Design**

The vital aspect of any program evaluation process is the selection of an appropriate and pertinent research design. Partly, the design of the study is determined based on the problems that should be investigated. Burns (2000, p. 3) defines research as “a systematic investigation to find answers to a problem.” More broadly, with regard to three approaches to research design (i.e. quantitative, qualitative, and mixed method), the factors that influence our choice include: “the research problem, the personal experiences of the researcher and the audience” (Creswell, 2003, p. 21). On the whole, language courses are dynamic

entities and subject to change according to the ever changing needs and wants of the students. To this end, Alderson (1996, p. 292) holds that “designs for evaluation should be practical, realistic and flexible.”

Obviously, the determining factor in the choice of one design over another is the purpose of the inquiry. “The potential range of methods is large, but the important element here is to ensure that they are accessible and reproducible; selection of them will depend on purpose and point of application” (Murphy, 1985, p. 11). On the other hand, Brown (1989) emphasizes that selecting the research design depends on the evaluation dimensions. That is, whether our evaluation is summative and carried out at the end of the program or formative and performed during the lifetime of the course. Also, whether our emphasis is on the product and/or what the students have achieved or on the process of teaching-learning process. Also, it should be clarified whether the evaluation emphasis is more on the quantitative or qualitative data. It can be added that the evaluation emphasis might be on both types of data at the same time. At this juncture, Blaxter *et al.* (2006) argue that if we could try to think methodologically on research design, the following questions might help illuminate our selection better. They (*ibid.* 62) argue that:

- What are the main purposes of your research?
- What is your role in the research?
- Do you think your values should affect research?
- Can you accommodate several methodologies in your research?
- Who are the audience for your research?

Therefore, it can be concluded that factors affecting the choice of one approach over another include the objectives, researcher’s experiences, nature of the problem(s), audience, availability of the resources and facilities, availability of the budget, and so on. Opting for a naturalistic and/or qualitative approach or a positivistic and/or experimental approach, Lynch (1996, p. 171) stresses “the inherent difficulties of trying to carry out experimental research in field settings, and of the limitations to what such an evaluation strategy can tell us about the programs we are investigating ...” He believes that applied

linguistics is gradually distancing itself from the experimental methods and moving towards the naturalistic inquiries. Burns (2000, p. 3) also is of the opinion that since 1960s the gulf between “the scientific empirical tradition and the naturalistic phenomenological method” has deepened. However, she (ibid.) prefers a naturalistic approach and argues that “Social reality is regarded as a creation of individual consciousness, with meaning and the evaluation of events seen as a personal and subjective construction.”

By and large, most of the researchers prefer a mixed method (i.e. both quantitative and qualitative approaches) because of its advantages and flexibility. Lynch (1996, p. 171) states that “The preferred evaluation approach ... is mixed strategies (i.e. quantitative analysis of qualitative data) or mixed design (positivistic and naturalistic).” The mixed method easily allows the inquirer to collect both numerical data and text data. The evaluator can use both types of data to increase the validity and reliability of his/her investigation. Creswell (2003, p. 22) also believes that “A mixed methods design is useful to capture the best of both quantitative and qualitative approaches.” He maintains that in order to gain a better understanding of the problem “both closed-ended quantitative data and open-ended qualitative data prove advantages” (ibid.). Mainly, the research design which best represents both characteristics of the qualitative and quantitative paradigm is the *illuminative model* of Parlett and Hamilton (1976). It is believed that this model suits the current exploration because of its appropriateness and adaptability. Therefore, it is elaborated on in the next section.

### **The Illuminative Model**

In the last decades there has been a wide range of program evaluation models available for the evaluators to select from and conduct their evaluation. One of the most prominent program evaluation models within the naturalistic design is the *illuminative model* developed by Parlett and Hamilton (1976). One of the significant advantages of this model is its utilization of both qualitative and quantitative data and analysis. For this reason Lynch (1996, p. 83) vehemently affirms that “illumination model is a mixed design rather than a naturalistic one.” He also argues that this model is at times called as transactional as a

result of “its focus on multiple audience perspectives and program process” (ibid. 82). Parlett and Hamilton (1976, p. 144) clarify their model’s aims as follows:

The aims of illuminative evaluation are to study the innovatory program: how it operates; how it is influenced by the various school situations in which it is applied; what those directly concerned regard as its advantages and disadvantages; and how students’ intellectual tasks and academic experiences are most affected.

It can be argued that illuminative model might be considered as one of the vivid examples of a nontraditional procedure for exploring the success or failure of a program. Richards (1984) elucidates illuminative model as a custom-built research strategy which generally keeps away from (though not ruling out) statistical procedures. This model attempts to obtain a wide variety of data on the program and its milieu and consequently work out its complexities as far as possible. This model’s advocates emphasize the importance of the process over the product (Ramsay & Clark, 1990). Also, one salient feature of the illuminative model is that it attempts to investigate all the aspects of the program and find out answers to different questions that may arise during the evaluation process. Richards and Schmidt (2002, p. 247) describe this model as:

an approach to evaluation that seeks to find out how different aspects of a course work or how a course is being implemented and the teaching-learning and processes that it creates. It seeks to provide a deeper understanding of the processes of teaching and learning that occur in a program...

There are usually three stages in which illuminative model can be implemented:

- 1- Observation,
- 2- Further inquiry and narrowing down the information,
- 3- Description and explanation of the findings.

In the first phase, the evaluator tries to obtain a holistic picture of the program's overall workings. He/She attempts to become acquainted with every aspect of the program and its context. After this initial period of observation, the evaluator makes every effort to derive minute issues and themes in order to narrow his/her studies and focus on the details. Parlett and Hamilton (1976) consider the movement from more general events to detailed ones as the progress focusing. Finally, the evaluator combines the findings and explains them to the intended audiences. One of the salient features of this model is that it deals with the unexpected happenings which may surface during the evaluation process. However, there is hardly any rigid borderline between these stages. "In practice, these stages overlap both temporally and functionally, and go on and on until, after successive inquiries, a clear focus is obtained" (Beretta, 1996, p. 17).

There are several major data collection techniques within the illuminative model: observations, interviews, questionnaires, program documentation, and so forth. By and large, observation is of prime importance in the initial stages of the evaluation. Then, the evaluator interviews program administrators, teachers and students. In order to increase the validity of the data, the evaluator gathers more information through questionnaires which are filled out by students, teachers and administrators. Moreover, the program's history, development and objectives can be investigated through program documentations.

Finally, the illuminative model has received some criticism now and then. For instance, Crittenden (1978, cited in Ramsey & Clark, 1990, p. 38) claims that:

First, by stressing the uniqueness of each setting, illuminative models do not produce findings which have any generalizability. Second, relying on the perceptions of the observer introduces problems of subjectivity. Third, the desirability, or even possibility of the evaluator remaining judgment free is questionable.

These criticisms are barely acceptable because the illuminative model has many advantages which make it almost one of the best methods of evaluating a program. First, it is an accepted fact that every educational setting has its peculiar characteristics. However, there are

also many similarities which cannot be ignored so easily. Furthermore, the aim of conducting an evaluation within a particular context is to find out its strengths and weaknesses. Therefore, the evaluation results are used to tackle that specific setting's problems and not to generalize them, though some generalizations can also be made to similar settings. Second, it is a myth that the findings are acquired merely as a result of the evaluator's perceptions. On the contrary, most of the findings are obtained based on program's stakeholders such as students, teachers and administrators through interviews and questionnaires. Finally, though the ultimate decisions are made by administrators and program organizers, it is the evaluator who observes, narrows the study and synthesizes the findings. It is, in fact, the evaluator who combines the various data and prepares the final report. On the whole, the illuminative model has many benefits which make it one of the prominent forms of program evaluation.

### **Conclusion**

In the main, evaluation is a multifaceted area in which various factors and elements should be taken into account. Program evaluation is not a simple one-shot procedure. It consists of different dimensions that determine the way in which varied types of information are collected. Also, program and/or course evaluation is carried out to achieve fairly different goals. Furthermore, program evaluation needs to be implemented through an appropriate and robust design. The design of the evaluation is of outmost importance. It is the design of the study which determines what type of data to be collected and how it should be analyzed and interpreted. Of course, the selection of a proper approach depends on the goals and objectives of the evaluation. In this process, the role of the evaluator is very critical. It is the evaluator who should opt for an adequate research design, collect relevant data, analyze them appropriately, and obtain the necessary results. All these efforts are made in order to produce an effective course and hence to modify and improve the learning-teaching process.

### References

- Afros, E. & Schryer, C. F. (2009). The genre of syllabus in higher education. *Journal of English for Academic Purposes*, 8, 224-233.
- Akst, G. & Hecht, M. (1980). Program evaluation. In A. S. Trillin, (Ed.), *Teaching basic skills in college* (pp. 261-310). San Francisco: Jossey-Bass.
- Alderson, J. C. (1996). Guidelines for the evaluation of language education. In J. C. Alderson & A. Beretta (Eds.), *Evaluating second language education* (pp. 274–304). Cambridge: CUP.
- Alderson, J. C. & Scott, M. (1996). Insiders, outsiders and participatory evaluation. In J. C. Alderson & A. Beretta (Eds.), *Evaluating second language education* (pp. 25–60). Cambridge: CUP.
- Beretta, A. (1996). Evaluation of language education: An overview. In J. C. Alderson & A. Beretta (Eds.), *Evaluating second language education* (pp. 5–24). Cambridge: CUP.
- Blaxter, L., Hughes, C. & Tight, M. (2006). *How to research*. Berkshire: Open University Press.
- 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: CUP.
- Brown, J. D. (1995). *The elements of language curriculum: A systematic approach to program development*. Boston, Mass.: Heinle & Heinle Publishers.
- Burns, R. (2000). *Introduction to research methods*. London: Sage Publications.
- Creswell, J. W. (2003). *Research design: Qualitative, quantitative, and mixed methods approaches*. Second ed. Thousands Oaks, CA: Sage Publications.
- Dudley-Evans, T. & St John, M. J. (2000). *Developments in English for specific purposes: A multidisciplinary approach*. Cambridge: CUP.
- Elley, W. B. (1989). Tailoring the evaluation to fit the context. In R. K. Johnson (Ed.), *The second language curriculum* (pp. 270–285). Cambridge: CUP.
- Jordan, R. R. (1997). *English for academic purposes: A guide and resource book for teachers*. Cambridge: CUP.
- Lynch, B. K. (1996b). *Language program evaluation: Theory and practice*. Cambridge: CUP.

- Murphy, D. F. (1985). Evaluation in language teaching: Assessment, accountability and awareness. In J. C. Alderson (Ed.), *Evaluation* (pp. 1–17). Oxford: Pergamon Press.
- Nunan, D. (1999). *Research methods in language learning*. Eighth printing. Cambridge: CUP.
- Parlett, M. & Hamilton, D. (1976). Evaluation as illumination: A new approach to the study of innovatory programs. In G. V. Glass (Ed.), *Evaluation studies* (pp. 140–157). Beverly Hills, Calif.: Sage Publications.
- Ramsay, W. & Clark, E. E. (1990). *New ideas for effective school improvement: Vision, social capital evaluation*. Hampshire: The Falmer Press.
- Richards, J. C. (1984). Language curriculum development. *RELC Journal*, 15(1), 1–14.
- Richards, J.C. & Schmidt, R. (2002). *Longman dictionary of language teaching and applied linguistics*. Third ed. London: Longman.
- Robinson, P. C. (1991). *ESP today: A practitioner's guide*. New York: Prentice Hall.
- Scriven, M. (1967). The methodology of evaluation. In R. W. Tyler, R. M. Gagne & M. Scriven (Eds.), *Perspectives of curriculum evaluation* (pp. 39–83). Chicago: Rand McNally.
- Sharp, A. (1990). Staff/Student participation in course evaluation: A procedure for improving course design. *ELT Journal*, 44(2), 132–137.
- Slimani, A. (1996). Evaluation of classroom interaction. In J. C. Alderson & A. Beretta (Eds.), *Evaluating second language education* (pp. 197–220). Cambridge: CUP.
- Tyler, R. (1949). *Basic principles of curriculum and instruction*. New York: Harcourt Brace.