

CHAPTER I INTRODUCTION

1.1 Background of the Study

For almost four decades there has been a strong tendency in applied linguistics in general and in program evaluation in particular to favor a traditional, quantitative experimental approach to conducting an evaluation study over a holistic, qualitative one. It was only in the past decade that a concern for tightly controlled experiments focusing on the analysis of product or student achievements has shifted to describing and analyzing the process of a program (Lynch, 1996).

The qualitative-quantitative debate has produced an alternative approach that is a combination of methods from the two paradigms: the positivistic and naturalistic approaches. Recently, there has been a growing number of evaluators who advocate evaluation approaches that combine both quantitative and qualitative methods (Lynch, 1996; Chen, 1997; Riggins, 1997; Smith, 1997). All methods have strengths and weaknesses. The combination of methods, or the so called "mixed-method design" can compensate for each method's weakness, triangulate the evaluative evidence, and expand the scope of study (Chen, 1997). It is worth noting that both quantitative and qualitative methods have their own criteria and merits, and qualitative methods cannot replace quantitative method, or vice versa. Mixed methods have own merits and are unlikely to replace other methods. On the other hand, they can expand options for evaluation design and data collection.

In fact, various approaches can be used to satisfy the purposes of different educational programs. No single approach can fit every program. In designing an appropriate evaluation, evaluators must consider the purposes of evaluation, and then select the most effective approaches that can be used in their study. For these reasons, a number of evaluation models have been developed and proposed by curriculum developers as well as evaluators in this field.

In Thailand, there is no single mandatory method of program evaluation, particularly at the tertiary level. Different institutions have different assessment schemes and criteria consistent with their national schemes, standards and educational policies. It seems that the most favourable forms of evaluation are questionnaires, and standardized proficiency and achievement tests. There is little evidence that indicates the use of naturalistic-qualitative methods such as action research. This may be

because most Thai evaluators have more concerns for investigating 'learning outcomes' which they think are more measurable and reliable than investigating 'learning processes'.

In most Thai universities, especially those that favour structural syllabuses, the students seem to have little confidence in using English and low communicative competence for the production of English. For these reasons, a task-based syllabus that can encourage maximal purposeful use of English directed towards learning and address students' apparent real world needs has become of interest to some avant-garde universities, such as King Mongkut's University of Technology Thonburi (KMUTT).

1.2 Statement of the Problem

A task-based EFL curriculum has been developed and used at KMUTT for more than seven years. Based on a learner-centered approach, it has its own unique characteristics. For example, it includes three types of analytic syllabuses: task-based, project-based and content-based syllabuses. The task-based syllabuses focus on the process of learning and on the students' needs in order to analyze the language input needed to understand specific language points. Each course is organized around goal-directed activities and the language to be learnt emerges from such activities (Watson Todd, 2001a). In the process of curriculum development, two main groups of people were involved: (1) the change agents, i.e. the senior members of the faculty, Assistant Professor Nuantip Tantisawetrat, the former Dean as well as Associate Professor Dr. Richard Watson Todd, a language specialist and research mentor who was the key person in curriculum innovation at KMUTT; and (2) the implementers, i.e. the staff in the Department of Language Studies that is responsible for English support courses at KMUTT (Watson Todd, R. 2001b). The task-based curriculum was proposed by the change agents who also largely made bureaucratic decisions about the proposed curriculum, but the majority of the work e.g., creating course outlines, tasks and materials was later completed mainly by the implementers with some input and feedback from the change agents. The process of curriculum renewal, then became an immanent process. The evaluations initially done in 2001 using action research (Kongchan, 2001; Wiriyaakrun, 2001) as well as the teachers' and students' informal reactions suggested that the first two courses that were designed, LNG 101 and LNG 102, together with the teaching materials seemed beneficial but not yet perfect.

Since 2001, the task-based curriculum has been informally revised almost every semester at the staff meetings for curriculum evaluation and modified many times by the program staff due to their personal beliefs that the curriculum itself contains a lot of drawbacks. This is called 'immanent curriculum renewal'. It refers to the involvement of teachers, as curriculum users, in the process of curriculum renewal. This teacher-based approach puts heavy emphasis on the process of development rather than the product. The curriculum is locally developed, information is gathered through action research, and the finished curriculum is expected to be adapted by the implementers as they see fit (Reid et al., 1987 cited in Watson Todd, 2006).

Watson Todd (2006) conducted his research to investigate why and how the curricula remained in a stage of immanent renewal, where the implementers felt they were in control of and free to make continual adjustments. The data for this study came from documentation concerning the course under investigation, and from interviews with teachers who had taught the courses regularly. The following table contains details of the course content and evaluation procedures of LNG over the four years that the study had been conducted (2001-2004). It also shows the revisions made each year.

Table 1.1: Summary of course changes for LNG102

LESSON	2001	2002	2003	2004
1	Introduction	Intro <i>Portfolio</i>	Intro <i>Portfolio</i>	Introduction
2	Self-access	Listening speaking	Listening speaking	<i>Portfolio</i>
3	Good lang. Learner	Listening speaking	Listening speaking	<u>Concordance dict.</u>
4	Learning strategies	Listening speaking	Listening speaking	<u>Concordance dict.</u>
5	Intro to skills	Listening speaking	Listening speaking	<u>Concordance dict.</u>
6	Intro to skills	<i>Portfolio consult</i>	<i>Portfolio consult</i>	<u>Concordance dict.</u>
7	<i>Portfolio</i>	<i>Portfolio consult</i>	<i>Portfolio consult</i>	<i>Portfolio consult</i>
8	Lang. Feedback	<u>Dictionary</u>	<u>Dictionary</u>	<i>Portfolio consult</i>
9	Resourcing	<u>Dictionary</u>	<u>Dictionary</u>	Speaking
10	Resourcing	Resourcing	<u>Concordance</u>	Speaking
11	Resourcing	Resourcing	<u>Concordance</u>	Resourcing
12	Resourcing	Resourcing	<u>Concordance</u>	Resourcing
13	Resourcing	Resourcing	Resourcing	Resourcing
14	Resourcing	Resourcing	Resourcing	<i>Portfolio consult</i>

LESSON	2001	2002	2003	2004
15	<i>Portfolio</i>	Resourcing	Resourcing	<i>Portfolio consult</i>
16	Lang. Feedback	<i>Portfolio consult</i>	Resourcing	<i>Portfolio feedback</i>
17	<u>Concordance dict.</u>	<i>Portfolio consult</i>	<i>Portfolio consult</i>	Resourcing
18	<u>Concordance dict.</u>	Resourcing	<i>Portfolio consult</i>	Resourcing
19	<u>Concordance dict.</u>	Resourcing	Resourcing	Resourcing
20	<u>Concordance dict.</u>	<u>Concordance</u>	Resourcing	Resourcing
21	Lang. Feedback	<u>Concordance</u>	Resourcing	Resourcing
22	Lang. Feedback	<u>Concordance</u>	Resourcing	Resourcing
23	<i>Problem solving</i>	Presentation	Presentation	Presentation
24	<i>Problem solving</i>	Presentation	Presentation	Presentation
25	<i>Problem solving</i>	<i>Portfolio consult</i>	<i>Portfolio consult</i>	<i>Portfolio consult</i>
26	<i>Problem solving</i>	<i>Portfolio consult</i>	<i>Portfolio consult</i>	<i>Portfolio consult</i>
27	<i>Portfolio</i>	Presentation	Presentation	Presentation
28	Lang. Feedback	Presentation	Presentation	Presentation
29	Summary	Summary	Summary	Summary
30	Summary	Summary	Summary	Summary
Evaluation procedures				
% for continuous assessment	75	65	65	60
% for exams	25	35	35	40

Note: *Lang.* stands for *Language*
Dict stands for *Dictionary*

Source: Watson Todd, R. (2006). Continuing change after innovation. *System*, 34, pp.1-14.

As shown in Table 1.1, in 2001, the original three main tasks in LNG 102 were resourcing, concordancing and dictionary work, and problem solving. In the introductory part, lessons were about the training of learning strategies and skills. A portfolio task was also employed as a course adjunct which ran through the course and several lessons were devoted to consultations, i.e. giving feedback on linguistic errors that students had made in working on the tasks and the portfolio. During 2002 to 2003, the course content as well as the assessment criteria had slightly changed. There were no more introductory lessons on skills. The problem-solving task was also eliminated.

Listening and speaking strategies were being taught in the same lesson. The scores for ongoing assessment was reduced from 75% to 65%, and the remaining 10% was added to the score on final exam. In 2004, the concordance task was finally deleted from the dictionary and concordance task. The score for the final exam was increased to 40%.

The findings from the interviews revealed the reasons for the changes, among them were the teachers' concerns about the lack of reliability in terms of assessment and evaluation procedures, incompatibility of students' needs with the course objectives— both linguistic and non-linguistic— and the impracticality of the learning and teaching methodology. As a result, attempts to increase the explicit teaching of linguistic objectives and to increase assessment through exams have been made to alter the original version of the task-based curriculum at King Mongkut's University of Technology Thonburi to the 'modified' one which lacks the features of no prespecified linguistic objectives and an emphasis on continuous assessment. These changes are still being questioned by the stakeholders and others on whether they could 'improve' or 'fail' the curriculum. As long as no more concrete and reliable evidence of the curriculum's ineffectiveness has been formally reported, i.e. no one can prove how effective it is, the curriculum is changed continuously. Watson Todd (2006)'s research reported only teachers' perspectives on the task-based curriculum and its process of changes. In fact, other stakeholders' voices should not be neglected. He, therefore, recommended that any appropriate system to guide any changes should be set up and implemented to determine the effectiveness of the innovation. Designing clearly targeted methods of obtaining students' and teachers' reactions to a new curriculum should be considered an essential part of the curriculum design process, and procedures need to be set up for putting these at the centre of any discussions to revise the curriculum.

Unfortunately, no single current evaluation model would fit the learner-centered task-based learning context. As this type of curriculum has its own characteristics and concepts different from any other kinds of curriculum, it needs a specific evaluation model that can fit its ideology and context.

In the mid-1980s, Beretta conducted an evaluation of a task-based English language project in Bangalore, India. The project was generally known as the Bangalore project, or the Communicational Teaching Project (CTP) which ran from 1979 to 1984. The Bangalore project was based on the generalization that the process of how learners learn a language structure is unconscious (Alderson and Beretta,

1992). The evaluation employed the large scale, product-oriented approach. A number of tests were administered to the CTP students: two achievement tests (a structure-based test and a task-based test); three proficiency measures; dictation, contextualised grammar; and a communicative listening/reading test. Obviously, Beretta's evaluation focused solely on student performance on tests. It could not explain what really happened in the classroom and how teachers implemented the program in the 'real' circumstances. Although Beretta later reported a retrospective study of the implementation of the CTP in the second phase of his study, only teachers were assigned to keep retrospective accounts without regard for students who are the key persons in learner-centered curriculum. It was suggested that in a task-based curriculum, students should not be treated only as testees; on the other hand, they need to participate actively in every step of the evaluation process.

In addition, some research studies into task-based learning reported that task-based approaches to instruction must place an emphasis on form without losing the value of tasks as realistic motivators, and as opportunities to trigger acquisitional processes (Skehan, 1996). Two models for balancing concern for 'communication' and 'form' are proposed by Willis (1996) and Skehan, (1996). The first model emphasizes a methodology for using tasks to combine natural communication with opportunities to focus on form. The second also shows how to balance form and meaning. Interestingly, such research studies and models reveal potential for organizing communicatively-oriented instruction which maintains a good balance between a concern for form and a concern for meaning. The modified task-based English curriculum at King Mongkut's University of Technology Thonburi may have been influenced by these approaches.

As no clear system or procedures for revising the curriculum have been set up to judge the effectiveness of this innovation since it was used for the first time in 2001, this research needs to be carried out. Due to time constraints, instead of examining the whole large-scale program, this study evaluates only a task-based English course at KMUTT by the use of a proposed model called the Integrated Stake-Tyler (or IST) Model developed by the researcher of this study. LNG 102, a task-based EFL course for undergraduate students at KMUTT, was employed as a testing ground.

In designing the proposed model, two preliminary steps are taken into account to find out the most appropriate evaluation for the given course: defining the function of the evaluation (formative or summative) and determining the approach to be used (a

single approach, or a combination of approaches) (Wilde and Sockey, 1995). Fitzpatrick, Sanders and Worthen (2004: 232) suggest some sources of information that can provide the foundation of the evaluation study: (1) questions, concerns, and values of stakeholders; (2) the use of evaluation models, frameworks, and approaches as heuristics; (3) models, findings, or salient issues raised in the literature in the field of the program; (4) guidelines developed or used elsewhere; and (5) the evaluator's own professional judgment. In formulating the IST model, the researcher searches information from these five different sources.

After having informal talks with the stakeholders as well as the audiences, the researcher finds that the issues they want the researcher to examine are student achievement and autonomy in language learning, and the needs of students and other stakeholders. The stakeholders and the audiences also want to know how effective the evaluated course is. Then, the researcher reviews literature on program evaluation to find out the model that best fit the purpose of the evaluation, and then finds that Stake's responsive approach is the most appropriate as it can evaluate students' learning outcomes, both product and process. Therefore, the researcher of this study has made a decision that to examine the effectiveness of the curriculum is the purpose of the evaluation. Both formative and summative evaluations are equally important and should be included in the same study. After having reviewed literature on program evaluation, the researcher finds that Stake's responsive model has limitation since it uses case study methods that may provide very subjective data. Tyler's objectives-based approach, which can also measure students' learning outcomes but provide more objective data, is included in this study to increase validity and reliability of the research. In developing evaluation questions, the researcher uses information from the existing program evaluation research and literature on applied linguistics. The issues found in Watson Todd's (2006) research, including lack of reliability in terms of assessment and evaluation procedures, incompatibility of students' both linguistic and non-linguistic needs with the course objectives, and the impracticality of learning and teaching methodology, have influenced the researcher's decision on the elements of the course. At this stage, three main features— context, implementation and student outcomes—that determine the effectiveness of the curriculum are identified. . Some guideline questions proposed by experts in program evaluation (Nunan, 1992; Brown, 1995; Lynch, 1996; Richards, 2001) reveal that the course elements that most evaluators examine are needs, goals and objectives, teaching, teacher, learner, course

content, resources, assessment and evaluation, materials, as well as stakeholders' opinions. Therefore, based on literature review and the researcher's own judgment, the researcher has decided to evaluate 12 dimensions: (1) needs, (2) goals and objectives, (3) teaching methods, (4) teachers, (5) tasks, (6) teaching materials, (7) resources, (8) assessment and evaluation, (9) students achievement, (10) students' autonomy in language learning, (11) students' opinions about the evaluated course, and (12) factors affecting students' learning outcomes.

In sum, the IST model is a mixed-method design that involves both qualitative and quantitative techniques in one evaluation since using multiple data-gathering techniques can provide all of the information needed for determining the effectiveness of the evaluated course.

1.3 Research Questions

Context

1. What is the context in which the curriculum is working?
 - 1.1 Do the goals and objectives of the course meet the needs of the stakeholders?
 - 1.2 Are the goals and objectives appropriate for the specified groups of students?

Implementation

2. How well is the curriculum being implemented?
 - 2.1 Is the teaching method relevant to the prespecified objectives?
 - 2.2 Are the teachers skillful in task-based instruction?
 - 2.3 Are the tasks related to the course objectives?
 - 2.4 Are the teaching materials relevant to the prespecified objectives?
 - 2.5 Are the resources adequate?
 - 2.6 Are the assessment procedures appropriate to the prespecified objectives?

Student outcomes

3. Are student outcomes due to the effects of what is happening within the curriculum
(context and implementation)?
 - 3.1 Do the students achieve the prespecified objectives?
 - 3.1.1 Do the students make significant gains in their language abilities after taking this course?

3.1.2 Do the students develop their autonomy in language learning during and after taking the course?

3.2 Do the learners think the course is appropriate?

3.3 Is the student learning the result of instruction or extraneous factors?

1.4 Objective of the study

The objective of this study is to evaluate the effectiveness of a task-based English course provided by the School of Liberal Arts at King Mongkut's University of Technology Thonburi based on a set of criteria and on the Integrated Stake-Tyler (IST) model especially proposed for this study.

1.5 Evaluation model and criteria

The IST model is based on Stake's responsive evaluation integrated with Tyler's objectives-based approach including some initiatives of the researcher of this study. The model includes active involvement of the stakeholders, students and program staff in particular, in the evaluation process; the multiple use of different measures, both quantitative and qualitative, in a single study; and use of portfolio assessment in course evaluation (See page 44 for more details).

The IST model is used to examine the effectiveness of the task-based course by examining three curriculum features: (1) context; (2) implementation; and (3) student outcomes, which can be subdivided into 12 dimensions. Each dimension is evaluated using different criteria, such as frequencies, a Chi-square test, an *F*-test and content-analysis (See Table 3.1 on pages 69-70 for more details).

1.6 Scope of the Study

This research was carried out to design an evaluation model appropriate for a task-based English course at King Mongkut's University of Technology Thonburi (KMUTT) which has its own distinctive characteristics:

1. The IST model would probably fit a task-based curriculum only. It may not be suitable for other types of curriculums, especially those that focus on 'product' rather than 'process'.
2. The subjects of this study include: (1) first- year KMUTT students who take a task-based EFL course, which is LNG 102; (2) English teachers who have experience in teaching the LNG 102 course; (3) subject teachers in the

- Faculties of Science, Engineering, and Industrial Education; and (4) audiences.
3. The scope of the evaluation is limited to the effectiveness of the evaluated curriculum. Course efficiency concerning investment of time and money is not included in the evaluation scheme.
 4. This evaluation investigates the effectiveness of the evaluated course by examining three main features: context, implementation, and student outcomes. They can be further divided into 12 dimensions: (1) needs, (2) goals and objectives, (3) teaching methods, (4) teachers, (5) tasks, (6) teaching materials, (7) resources, (8) assessment and evaluation, (9) students achievement, (10) students' autonomy in language learning, (11) students' opinions about the evaluated course, and (12) factors affecting students' learning outcomes.
 5. There are two English courses using the same title, LNG 102, one of which is the currently existing task-based course. The evaluated course is the existing task-based course that was taught in the second semester of the academic year 2006. Therefore, the results of this study cannot be generalized to the intensive one recently designed in June 2006 for a 12-day intensive program.

1.7 Limitations of the Study

1. The samples of this study are intact groups in the fields of science, industrial education and engineering; therefore, the results of the study may not be generalized beyond these groups.
2. Since the Integrated Stake-Tyler (IST) model was especially designed for this highly context-specific study, the findings from this study may not be generalized to other contexts.
3. The instruments, namely, curriculum-based achievement tests, student portfolios, semi-structured interviews, self-assessment checklist, and classroom observation, were carefully selected and designed to fit the purpose of this study. They may not be suitable for other evaluation studies.

1.8 Definition of Terms:

Audiences refer to people who will make use of the evaluation results. In this study, they refer to the Dean of the School of Liberal Arts and the Head of the Department of Language Studies who are interested in the results of the evaluation, and can make any judgement on the evaluated course, such as whether it should be retained, revised, or eliminated.

Current students refer to KMUTT students who took the LNG 102 course in the second semester of the academic year 2006.

Degrees of effectiveness refers to the extent of success of the evaluated course determining by comparing the overall scores obtained from the the evaluation of the evaluative dimensions with the grading criteria for undergraduate courses at Chulalongkorn university (CULI, 2007):

85%-100% = very successful

75%-84% = successful

65%-74% = fairly successful

55%-64% = partially successful

0%-54% = fail

Effectiveness refers to the success of the task-based English course at KMUTT that can be determined by analyzing three main features of the curriculum: context, implementation and student outcomes. Context includes needs, as well as goals and objectives. Implementation includes teaching methods, teachers, tasks, teaching materials, resources, as well as assessment and evaluation. Student outcomes include the students' achievements, autonomy in language learning, opinions about the evaluated course, and factors affecting their learning outcome. The evaluated course can be considered effective if these evaluative dimensions meet the criteria set specially for this evaluation.

Evaluated course refers to the LNG 102 that was taught in the second semester of the academic year 2006.

Evaluation model refers to a particular design for evaluating a certain type of curriculum.

Ex-students refer to students who have already taken the LNG 102 course.

Integrated model refers to the proposed evaluation model designed by the

researcher of the study. It is based on Stake's and Tyler's evaluation models. It is the mixing of naturalistic and positivistic designs at the methodological level. In this model, the data are analyzed qualitatively and quantitatively.

Portfolios refers to an adjunct activity which students undertake in parallel with the main tasks of the LNG 102 course. It aims to improve students' English proficiency and to enhance their awareness of autonomous learning and self-directed work.

Stakeholders refer to students, English teachers, and subject teachers from different faculties at KMUTT.

Subject teachers refer to KMUTT teachers who teach subjects other than English.

Task-based English course refers to LNG 102: Fundamental English II which is the second in a four-course series in the English as a Foreign Language curriculum using a task-based approach provided by the School of Liberal Arts, KMUTT. The important elements of the course are composed of: (1) concepts and goals; (2) tasks; (3) teaching materials; (4) resources; and (5) assessment.

Teachers refer to teachers of English at the Department of Language Studies, School of Liberal Arts, King Mongkut's University of Technology Thonburi who have experience in teaching a task-based English course -- LNG 102: Fundamental English II.

12-Dimension questions refer to the evaluation questions identified by the researcher of this study. They are developed from the 12-evaluative dimensions—the indicators of effectiveness of the evaluated course.

1.9 Significance of the Study

1. The information gained from this research will be evidence for the curriculum developers at KMUTT in making judgement on the quality of the evaluated task-based EFL course, i.e. what elements in the curriculum should be retained, improved or eliminated.

2. If this research yields positive results, the Department of Language Studies, School of Liberal Arts, King Mongkut's University of Technology Thonburi will be able to use the research results to promote the accountability of its EFL program.

3. If the research results proves that there are some flaws in any components of the evaluated course, the curriculum developers as well as the teachers can utilize the findings for the purpose of course improvement.

4. If this research proves that the proposed model can work effectively, it can be used as an evaluation model for other task-based EFL courses at KMUTT.

5. If this research proves that the proposed model can work effectively, it can be applied to the evaluation of any EFL task-based courses in more or less similar contexts.

6. If it is found that the proposed model is unworkable, there should be further research to find an evaluation model that can examine task-based English courses more effectively.

1.10 Overview of the Dissertation

The dissertation is composed of five chapters:

Chapter I begins with the background of the study followed by the statement of the problem, research questions, objectives, scope of the study, definition of terms as well as the significance of the study.

Chapter II presents the underlying theoretical framework including a task-based approach, self-directed learning and learner autonomy, curriculum (or program/course evaluation), the responsive approach, and the Tylerian approach. The Integrated Stake-Tyler model specially designed for the study is proposed. Research studies in program evaluation are also included.

Chapter III describes the research methodology of the study which includes the research procedures, population and samples, research instruments, research instrument construction and validation, the methods of data collection and data analysis, as well as statistical tools.

Chapter IV deals with the results of the study regarding the research questions.

Chapter V concludes the study by presenting the summary and discussions of the findings and recommendations for curriculum developer, teachers and researchers.