

## CHAPTER III

### RESEARCH METHODOLOGY

This chapter presents the details regarding the procedures and methodology of the research study. The main objectives of the study are to develop the Active Learning Instructional (ALI) Model for enhancing secondary school students' English communicative abilities in large classes, to evaluate the effectiveness of the ALI Model in a large Secondary-Education-Level authentic language classroom by the experts and through experimentation, and to study the students' opinions towards the ALI Model for enhancing their English communicative abilities in large classes.

#### Statement of Hypotheses

1. The English proficiency mean score of the students, who receive the treatment based on the ALI Model, is significantly higher than that of the students who are controlled to receive the conventional way of teaching. ( $H_1: \mu_1 > \mu_2$ )

2 The English performance mean score of the students, who receive the treatment based on the ALI Model, is significantly higher than that of the students who are controlled to receive the conventional way of teaching. ( $H_2: \mu_3 > \mu_4$ )

3 The students in the treatment group are likely to have positive opinions towards the ALI Model for enhancing secondary school students' English communicative abilities in large classes. ( $H_3$ : Mean of opinion scale  $\geq 3.5$  from the 5-point scale on the questionnaire)

#### Research Design

This study is developmental and experimental research. The research has been conducted in two main phases as follows:

Phase 1: The development of the Active Learning Instructional (ALI) Model for enhancing secondary school students' English communicative abilities in large classes. There were four stages in this phase:

Stage 1: Studying, analyzing, and synthesizing the theoretical and pedagogical principles for teaching English communicative abilities in large classes via active learning

Stage 2: Developing the ALI Model rationales

Stage 3: Determining the ALI Model framework

Stage 4: Developing the ALI Model steps

Phase 2: The evaluation of the effectiveness of the Active Learning Instructional (ALI) Model for enhancing secondary school students' English communicative abilities in a large class. There were two main stages:

Stage 1: Validating the ALI Model by experts

Stage 2: Implementing the ALI Model in a large authentic classroom

To validate the ALI Model, three experts in the field of ESL/EFL were asked to verify the model using the evaluation form. In order to evaluate the effectiveness of the ALI model, quantitative data was mainly considered. Qualitative data, nevertheless, was used to confirm the results and provide insights into the study. This is because the quantitative data alone may not be enough to justify the trustworthiness of the study. However, qualitative data can provide in-depth information which can add insights into the understanding of the phenomenon that might be missed when only one method is employed (Onwuegbuzie & Leech, 2004).

The methodology of evaluating the effectiveness of the ALI Model was based on the *quasi-experimental design* (Issac & Michael, 1981) comparing English communicative abilities of students in the experimental group—obtaining the treatment of ALI Model instruction—and those in the control group—receiving the treatment of the PPP Method—via a proficiency test and a performance test. The design in Table 3.1 indicates that T represents the tests, while X represents the teaching methods.

Table 3.1: Representation of a research design

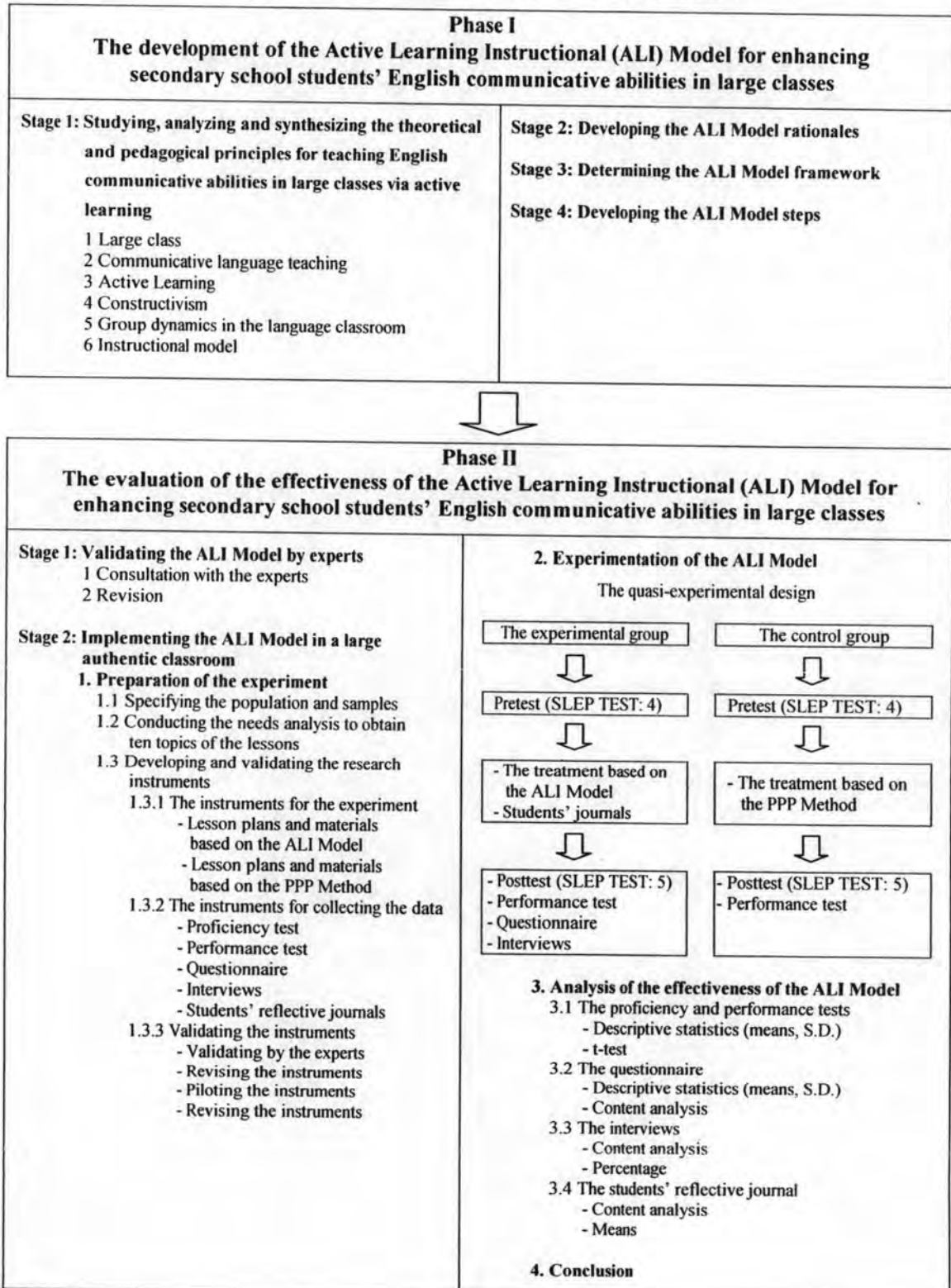
Group A: Experimental Group	$X_1$	$T$
Group B: Control Group	$X_2$	$T$

In this case, the  $X$ s were independent variables referring to the teaching modes used in this study. The  $X_1$  represented the ALI Model instruction whereas the  $X_2$  referred to the PPP Method. The  $T$  represented the posttest. The students' English communicative ability scores on this measure were dependent variables of the study. In the experiment, the contents and lessons were similar in both groups. The difference was found in the modes implemented in this study, which served as the independent variable. A measure to assess students' English communicative abilities was a parallel form consisting of a SLEP Test and a performance test. Students' test scores were used to estimate the statistical significance of between group differences.

To investigate students' opinions towards the ALI model instruction, students' reflective journals, questionnaires, and interviews were employed.

Two phases of the research procedures were summarized and are presented in the following figure.

Figure 3.1: The summary of research procedures



**Phase 1: The development of the Active Learning Instructional (ALI) Model for enhancing secondary school students' English communicative abilities in large classes**

Stage 1: Studying, analyzing and synthesizing the theoretical and pedagogical principles for teaching English communicative abilities in large classes via active learning

The researcher studied the basic knowledge from various books, journals, websites and related research consisting of the following information:

1. Information on current issues about class size of language classrooms as well as teaching and learning communicative English

To ascertain whether class size was really a main problem and was significant enough to be worth studying, a background study about experience and perception of class size (Appendix A) was performed with teachers who taught English in state secondary schools in Bangkok. The overall findings of the study were satisfactory and confirmed the possibility and value of doing the project.

2. Information on the school curriculum and the national curriculum of the English subject for the secondary educational level in terms of objectives, content, pedagogical activities and assessment and evaluation

3. Information on theories and concepts relevant to the ALI Model, including large class, communicative language teaching, active learning, constructivism, group dynamics in the language classroom and instructional model

After studying the aforementioned concepts and principles, the key concepts of the study were analyzed and synthesized.

Stage 2: Developing the ALI Model rationales

Based on the key concepts obtained from the study, analysis and synthesis of the theoretical and pedagogical principles for teaching English communicative abilities in large classes via active learning, the ALI Model rationales were developed accordingly.

Stage 3: Determining the ALI Model framework

The ALI Model framework, which consisted of model rationales, objective, model steps and outcomes, was determined.

Stage 4: Developing the ALI Model steps

From the model rationales synthesized from the theoretical and pedagogical principles, the ALI Model steps were developed for enhancing secondary school students' English communicative abilities in large classes.

**Phase 2: The evaluation of the effectiveness of the Active Learning Instructional (ALI) Model for enhancing secondary school students' English communicative abilities in a large class**

Stage 1: Validating the ALI Model by experts

1. Consultation with the experts: Three experts in ESL/EFL were asked to verify the model by providing each of them with an evaluation form (Appendix C). There were two parts on the form: the four-point attitude scale and an open-ended question. There were five items in the first part of the evaluation form asking the experts to mark each item as appropriate (+1), not sure (0), or not appropriate (-1). The other part was an open-ended question asking the experts to show their additional comments and recommendations. After that, the experts' responses from the first part were calculated by means of Item-Objective Congruence Index (IOC) (Turner & Carlson, 2003).

$$IOC = \frac{R}{N}$$

IOC	means	the index of congruence
R	means	total score from the opinion of the specialists
N	means	numbers of the specialists

If the IOC value is higher than 0.5, it is accepted. However, if it is lower than 0.5, that point must be revised. This approach helped confirm the reliability and validity of the model. The experts also provided their additional comments after verifying the model.

2. Revision: Considering comments obtained from the experts, some changes were made to improve the model.

The detailed information of the experts' validation is presented in chapter 4.

Stage 2: Implementing the ALI Model in a large authentic classroom

In this stage, there were four sub-stages: preparation of the experiment, experimentation of the ALI Model, analysis of the effectiveness of the ALI Model, and conclusion.

1. Preparation of the experiment

There were three steps for preparing the experiment: specifying the population and sampling, conducting a needs analysis, and developing and validating the research instruments.

1.1 Specify the population and sampling

### 1.1.1 Population

The population of this study included 240 Grade-10 students who were studying Foundation English 2 in semester two, academic year 2005 at a secondary school in Nakhon Pathom Province, Thailand. The students from this school were selected because the director of the school agrees that large class size is a common problem in the school and would like to see how students can be motivated to learn in such classes. The research proposed the concept of the ALI Model to help enhance secondary school students' English communicative abilities in large classes via learning, doing, observing and reflecting. Through this process, the students would actively participate in class. Therefore, with the consent of the school director, the ALI model would serve as a significant tool in developing a Grade-10 English communicative course for the academic year 2006.

### 1.1.2 Subjects and Sample selection

According to the school admission, there is no English examination. The students are required to take only Science and Mathematics examinations. Therefore, students in each class at this school are multilevel in English proficiency. Every student must take an English course (2 credits) in every semester.

The selection of samples from the whole population was done by means of the purposive selection. To ensure that the two groups were comparable, group equality was statistically verified by comparing the pretest mean scores (the proficiency test: SLEP test form 4) using the *Independent samples test* of the SPSS program at a significance level of 0.05.

The pretest scores of both groups were not significantly different at the beginning of the study. It was found from the mean comparison that both groups were somewhat comparable in their English proficiency at the level of 0.05 as illustrated in Table 3.2.

Table 3.2: Mean comparison of pretest scores between the control and experimental groups

Group	n	$\bar{X}$	S.D.	t	P
Control	48	44.46	7.478	.580	.563
Experimental	48	43.65	6.183		

\*p<.05

Both groups, containing 48 students in each, were then randomly selected as being the experimental group receiving the treatment based on the ALI model and the control group receiving the instruction based on the PPP method.

### 1.2 Conducting a needs analysis

In this study, the needs analysis—the logical starting point for the development of a language program which is responsive to the learner and learning needs (Graves, 2000; Richards, 2001)—was conducted in semester one of academic year 2005 in order to obtain ten topics of the lessons. Fifty-six topics from five main areas in the theme generator wheel (Estaire & Zanon, 1994) as illustrated in Figure 3.2 were translated into Thai and presented to the subjects (96 students in the course “Foundation English 1” in semester one of academic year 2005). The subjects were asked to choose ten out of 56 topics that they were interested in (Table 3.3). Finally, the top ten selected topics served as topics for a ten-unit English communicative course.

Figure 3.2: The theme generator wheel (Estaire & Zanon, 1994)

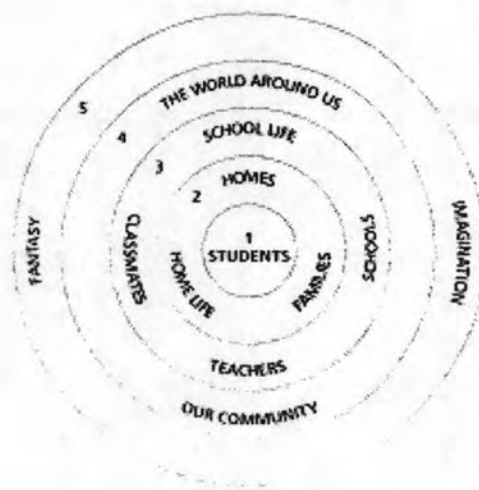


Table 3.3: Fifty-six topics from the theme generator wheel (Estaire &amp; Zanon, 1994)

<p><b>Circle 1</b></p> <p>1 Our birthdays 2 Our eating habits 3 Our body (limited to parts of the body) 4 How our body works 5 What will we be like in 20 years' time?</p> <p><b>Circle 2</b></p> <p>6 What were our parents like 20 years ago? 7 Our homes 8 The pocket money we get 9 Who does the housework at home: a survey and discussion 10 What time should we get back home? 11 Arguments at home 12 A survey of men's jobs and women's jobs in our families: conclusions</p> <p><b>Circle 3</b></p> <p>13 School rules (including responsibilities and rights) 14 Our ideal school 15 Exams 16 Preparing a party (e.g. carnival) at school 17 Our class/school magazine 18 Learning</p> <p><b>Circle 4</b></p> <p>19 Smoking 20 Nuclear power 21 Where people live 22 Space travel 23 Changes in everyday life from the time of our grandparents up to now 24 Planning a journey 25 Drugs 26 Water 27 Advertising 28 Countries (places) we would like to visit 29 Our ideal neighborhood/village/town 30 A news program</p>	<p>31 Traditions in our community/in English-speaking communities 32 Religions of the world 33 Health (including a class record of absences due to health problems) 34 Interviewing English-speaking tourists in our area 35 Letters to pen friends (or cassettes/videos/disks) 36 English-speaking people in our community 37 English around us 38 A writer/a painter: life and work 39 Animals in danger 40 National parks around the world 41 Our neighborhood or town: past, present, future 42 Man and nature 43 Hunger in the world 44 Christmas 45 Our language: origins, history 46 Describing people 47 Interesting people in the community 48 Gone missing 49 Go Green 50 People from other countries living in our community 51 Military service: opinions of Military service: a comparative study of EC countries 52 Being young in the '90s (or in the year 2000, when the time comes!)</p> <p><b>Circle 5</b></p> <p>53 Mystery 54 Dreams 55 A short play (writing and producing it, perhaps based on a story read) 56 A story (writing it)</p>
---	--

Ten topics from the needs analysis are presented in the following table according to the frequency.



Table 3.4: Frequency of the top ten selected topics obtained from the needs analysis

Topics	Frequency
1. Movies	78
2. Traveling	72
3. Cooking	68
4. Our birthday	60
5. Advertising	60
6. Mystery	53
7. Describing people	49
8. School Rules	44
9. Agony Aunt	43
10. Animals in danger	37

Ten topics obtained from the needs analysis were rearranged according to the five main areas in the theme generator wheel (Figure 3.2) from the innermost part of the wheel, which was most relevant to the students' background, to the outermost, which was least relevant to the students' background, and are presented in the following table.

Table 3.5: Ten topics rearranged according to the theme generator wheel

Lesson	Content
1	Our birthday (Our birthday: Circle 1)
2	Cooking (Our home: Circle 2)
3	Agony Aunt (Argument at home: Circle 2)
4	School rules (School rules: Circle 3)
5	Traveling (Planning a journey: Circle 4)
6	Advertising (Advertising: Circle 4)
7	Animals in danger (Animals in danger: Circle 4)
8	Describing people (Describing people: Circle 4)
9	Mystery (Mystery: Circle 5)
10	Movies (A short play: Circle 5)

### 1.3 Developing and validating the research instruments

There were two main types of research instruments in the study: the instruments for the experiment, and the instruments for collecting the data.

1.3.1 The instruments for the experiment include ten lessons plans based on the ALI Model and ten lesson plans based on the conventional way of teaching (PPP Method). Each lesson plan for both groups was designed for a 100-minute instruction in two periods, which was based on the normal practice of the school schedule.

#### 1.3.1.1 Lesson plans based on the ALI Model

Ten topics obtained from the needs analysis were designed to be ten lesson plans, each of which was based on the ALI Model (See Table 4.1), for teaching students in the experimental group. At the end of each lesson plan, students were asked to write a five-minute students' reflective journal. Each lesson plan (an example is attached in Appendix D) was designed according to the ALI Model steps.

#### 1.3.1.2 Lesson plans based on the PPP Method:

With the same topics, ten lesson plans were designed according to the PPP Method. PPP stands for Presentation (introduction of a new teaching item in context), Practice (controlled practice of the item), and Production (a free practice phase) (Harmer, 2001: 80). The lesson plans (an example is attached in Appendix E) were designed based on the following steps of the PPP Method.

Table 3.6: Designing lesson plans based on the PPP method

STEPS	STRATEGIES
1. Presentation	1. T presents Ss the word list and grammar points relevant to the topic and content.
2. Practice	1. T gives Ss a model exercise based on the vocabulary and grammar points introduced in step one and lets Ss practice by doing the exercise. The activity is based on individual work. <ul style="list-style-type: none"> <li>1.1 For reading tasks               <ul style="list-style-type: none"> <li>- T randomly asks Ss one by one to read one sentence in the reading text and to translate that into Thai. If there are any pronunciation or translation mistakes, T corrects them and lets Ss repeat the correct version again.</li> <li>- T lets Ss do the exercise and then gives them the correct answers.</li> </ul> </li> <li>1.2 For listening tasks               <ul style="list-style-type: none"> <li>- T lets Ss listen to the tape twice.</li> <li>- T lets Ss do the exercise after listening to the tape.</li> <li>- T gives the correct answers to Ss.</li> <li>- T lets Ss listen to the tape once again.</li> </ul> </li> </ul>
3. Production	1. T asks Ss to produce a piece of work according to the exercise in the <i>Practice</i> step. The activity can be pair or group work. 2. T asks Ss to present their work in front of the class. The activity can be pair or group work.

#### 1.3.1.3 Validating the lesson plans by the experts and revision

The researcher constructed a research instrument evaluation form (Appendix F) to evaluate the effectiveness of the lesson plans. The quality level of such was obtained by the submission of three lesson plans (Units 2, 5, and 10) as representatives from various main areas in the theme generator wheel to three experts

in the fields of ESL/EFL and teaching methodology. The research instrument evaluation form consisted of seven semi-structured questions and asked the experts to write their comments and suggestions as they pleased. The comments are summarized in the following table.

Table 3.7: Summary of the experts' comments toward both types of the lesson plans

Items	COMMENTS	
	The lesson plans based on the ALI Model (The experimental group)	The lesson plans based on the PPP Method (The control group)
1. Objectives of the lesson plans	<ul style="list-style-type: none"> <li>- There was some overlapping of some objectives.</li> <li>- The language should be more consistent.</li> </ul>	<ul style="list-style-type: none"> <li>- There was some overlapping of some objectives.</li> <li>- The language should be more consistent.</li> </ul>
2. Materials used in the lesson plans	<ul style="list-style-type: none"> <li>- Most of materials were teacher-made. Students should be free to select their own materials (under teacher guidance), analyze them, and come up with their own explanation.</li> </ul>	<ul style="list-style-type: none"> <li>- Some lessons had not enough materials, especially those dealing with language awareness. For example, in unit 2: Cooking; examples on the use of transitional words should be provided. Moreover, in unit 5: Traveling; descriptive phrases should be emphasized.</li> </ul>
3. Evaluation used in the lesson plans	<ul style="list-style-type: none"> <li>- Peer feedback should involve some discussion rather than just checking the answers with friends.</li> <li>- Some guidelines must be given for peer editing and peer revision.</li> </ul>	<ul style="list-style-type: none"> <li>- The rubric set for each task should be provided to students.</li> </ul>
4. Language used in the lesson plans	<ul style="list-style-type: none"> <li>- The language should be more consistent.</li> <li>- Two more columns (materials and evaluation) should be added in the format of the lesson plans.</li> <li>- There were some language redundancies.</li> </ul>	<ul style="list-style-type: none"> <li>- The language should be more consistent.</li> <li>- Two more columns (materials and evaluation) should be added in the format of the lesson plans.</li> <li>- There were some language redundancies.</li> </ul>
5. Pedagogical procedures in the lesson plans (e.g. are the activities interesting and fostering students' English communicative abilities?)	5.1 Warm-up <ul style="list-style-type: none"> <li>- More time might be needed to do all the activities.</li> </ul>	5.1 Presentation <ul style="list-style-type: none"> <li>- More language focus should be provided to make sure that students would be well-equipped before moving to the practice stage.</li> </ul>
	5.2 Individual Learning <ul style="list-style-type: none"> <li>- Some activities were traditional, teacher-directed.</li> </ul>	5.2 Practice <ul style="list-style-type: none"> <li>- Pair/Group work might be used where appropriate.</li> </ul>
	5.3 Small-Group Learning <ul style="list-style-type: none"> <li>- The composition of the group should be specified (e.g. random or mixed ability).</li> </ul>	5.3 Production <ul style="list-style-type: none"> <li>- It was quite unpractical to ask each of the students in a large class to give a presentation.</li> </ul>
	5.4 Large-Group Learning <ul style="list-style-type: none"> <li>- How large is a large group?</li> </ul>	
6. The clear difference of the lesson plans between experimental and control groups	<ul style="list-style-type: none"> <li>- The activities in step 1-3 in the experimental group were not much different from those in the control group.</li> <li>- I was wondering why pair/group work was not used at all in the control group.</li> <li>- Students in the experimental group should take charge of their own language more, think more globally (in relevant contexts), and use more analytical and critical skills rather than starting with discrete activities as proposed.</li> </ul>	
7. Additional comments/suggestions	<ul style="list-style-type: none"> <li>- The large-group activities were well selected as they were realistic tasks.</li> <li>- Student-teacher conferences should be included in the ALI Model lesson plans as a form of feedback.</li> <li>- Students in the experimental group should be trained to do peer revision.</li> </ul>	

The results from the evaluation form revealed that the experts all agreed with the lesson plans in terms of topics, objectives, procedures, and evaluations. However, there were some flaws in details that the experts suggested changing (Table 3.7). The researcher adjusted the lesson plans according to the experts' suggestions and then asked them to approve the revised versions. The final versions were approved by the experts.

#### 1.3.1.4 Pilot of the lesson plans and Revision

To ascertain that the actual two types of lesson plans were effective and to reduce unforeseen problems, a pilot study was carried out in semester one of academic year 2005. The samples in the pilot study, which had equivalent characteristics as the subjects in the main study, consisted of 96 Grade-10 students at this school who were divided into two equal groups (48 students for each group). One group received the instruction of three ALI Model lesson plans, and the other received a treatment of three lesson plans based on the PPP Method. The pilot study reduced the number of treatment errors, because unforeseen problems revealed in the pilot study were overcome in redesigning the main study. It was possible to get feedback from participants and other people involved which led to improvements such as in alternative instruments, materials, and procedures.

After the pilot, the main flaw that was found was timing. It obviously showed that some steps contained too many activities and took too much time. Therefore, some activities were deleted and some were mixed or combined into one. Moreover, the order of some activities was changed and rearranged in order to make the lesson run smoothly.

1.3.2 The instruments for collecting the data include proficiency test, performance test, Questionnaire for Eliciting Students' Opinions, interviews, and students' reflective journals.

##### 1.3.2.1 Proficiency Test

Two parallel forms (Forms 4 and 5) of the Secondary Level English Proficiency (SLEP) Test (See Appendix G) served as the pretest and posttest of the research respectively. According to the SLEP Test Manual (ETS, 2004: 5-6), the SLEP test is a test measuring English language ability and is designed for use with students entering grades seven through twelve or community colleges whose native language is one other than English.

The SLEP test was developed by staff at Educational Testing Service with the advice and assistance of a committee of examiners composed of secondary school ESL teachers with extensive experience teaching English to adolescent students. The test is divided into two sections, each containing four types of questions. For the first section, the four types of questions all use recorded samples of spoken English to test listening comprehension and do not rely heavily on written material. The four question types in the second section, Reading Comprehension, are based on written or visual materials. Section two includes written questions based on a cartoon, written questions based on line drawings, three multiple-choice cloze passages, and a literary passage followed by questions on its content. This section also measures vocabulary and grammar. Test questions in both sections of the SLEP test are based on information presented in or easily inferred from the questions or from the associated passages or pictures. Knowledge of specific subject matter is not tested, and there are no questions that have a bearing on literary knowledge, literary analysis, or linguistic terminology.

The reliability of each form of SLEP has been estimated using an internal-consistency measure of reliability, Cronbach's coefficient alpha. Reliabilities and standard errors of measurement (SEM) for each section of the test, as well as the total test, are provided below.

Table 3.8: Reliability and SEM of SLEP Test form 4 and 5

Form	Section One		Section Two		Total Test	
	Listening Comprehension		Reading Comprehension		Reliability	SEM
	Reliability	SEM	Reliability	SEM		
4	.94	1.5	.88	1.8	.95	2.3
5	.93	1.6	.91	1.4	.95	2.1

#### 1.3.2.2 Performance Test

Most educators agree that standardized, norm-referenced tests are inappropriate for English language learning students since traditional forms of assessment such as standardized tests use multiple choice items (O'Malley & Valdez Pierce, 1996: 3-4). Furthermore, multiple-choice items assume a level of English language proficiency that English language learning students may not have acquired. The subtle distinctions made on various items for vocabulary, word analysis, reading and listening subtests may produce information on what the student does not know but little information about what the student does know. This gives the teacher an

incomplete picture of student needs and strengths (Stiggins, 1991). The language components of standardized tests mainly assess reading and vocabulary knowledge and ignore progress in written and oral language, which are important components of language-based instruction programs (O'Malley & Valdez Pierce, 1996: 3-4). However, this does not suggest that there is no role for standardized testing in the assessment of English language learning students. Standardized tests have an important role in at least four components of an overall testing program: 1) to compare individual or group performance with an external normative group, 2) to identify relative strengths and weaknesses in skill areas, 3) to monitor annual growth in skills and 4) for program evaluation (Hoover, 1995 cited in O'Malley & Valdez Pierce, 1996: 4). Nevertheless, the use of standardized tests does not cover the full range of assessment needs for English language learning students; thus, alternative (or authentic) assessment comes into play (O'Malley & Valdez Pierce, 1996: 4).

Alternative (or authentic) assessment is described as the multiple forms of assessment that reflect student learning, achievement, motivation and attitudes on instructionally-relevant classroom activities. Examples of alternative assessment include performance assessment, portfolios, and student self assessment (O'Malley & Valdez Pierce, 1996: 4).

In this study, a performance test is another main instrument for collecting the data. Performance test is defined as a measure of assessment based on authentic tasks such as activities, exercises, or problems that require students to show what they can do (McBrien & Brandt, 1997: 77). The performance test is an accurate reflection of authentic language use for finding out what a student knows or can do that is intended to show growth and inform instruction, and is an alternative to traditional forms of testing, such as multiple-choice tests (Stiggins, 1991). In this study, a performance test (see Appendix H) serves as an alternative assessment which served as the achievement test, which was constructed to measure students' English language abilities via authentic tasks. The test was used to provide more accurate measures of progress towards communication proficiency goals. The test is criterion-referenced and is authentic because it is based on the individual and group-work activities that represent classroom and real-life settings (O'Malley & Pierce, 1996: 1-2). There were two parts of the test. First, the students were asked to select one favorite article of 150-300 words in length and then let them individually write a paragraph of 120-150 words showing three opinions toward the selected article. Then

the students' writing assignments were compiled as a class magazine. There were 20 points for this writing task. The other part of the test, containing 20 points, was the presentation. The rubrics set of writing test and presentation were adapted from the school's rubrics (Appendices I and J). Inter-rater (Brown, 2005) was employed to grade the students' performance test. For each student, the total score of each part of the performance test was 20. However, the final score of the performance test for each student was derived from the score of the writing part plus the score of the presentation, divided by two. Therefore, the finalized total performance test score for each student was 20.

Three experts in the fields of ESL/EFL teaching and language assessment were asked to validate the performance test and rubrics. They suggested that some redundancy of the language and some complicated parts should be modified. Moreover, they said that the steps of the test should be presented in accordance with the ALI Model steps. The revision was made according to the experts' suggestions. A trial of the revised performance test was made with a group of 48 Grade-10 students who participated in the trial of the ALI-Model lesson-plan verification in order to see existing or potential problems so that corrections and adjustments could be made before being used in the main study. The students were asked to do the writing and presentation tasks according to the set performance test. Two teachers graded the students' performance test in order to obtain the inter-rater reliability (Bachman, 2004: 169). The scores were then statistically verified to find out the correlation with the teachers' grading. The correlation of the scores between both raters, using the Pearson product-moment correlation coefficient (Bachman, 2004: 85), was 0.803 at the significance level of 0.01. Therefore, this showed that both teachers had significant relationship of rating. In the main study, there was one more criterion; that for each student, the scores gained from both teachers must not be more than 3 points ( $\pm 3$ ), then the real score of a student would be derived from teacher A's score plus teacher B's score divided by two. However, if any scores from both teachers were more different than 3 points, discussion would be made for the rating revision. After the trial, there was no flaw in the performance test, so it was ready to be used in the main study.

### 1.3.2.3 Questionnaire for Eliciting Students' Opinions

The Questionnaire for Eliciting Students' Opinions was a 5-point Likert scale questionnaire. The five options were 1) strongly disagree, 2) disagree, 3)

undecided, 4) agree and 5) strongly agree. The questionnaire was designed to collect students' demographic characteristics and their opinions towards the instruction based on the ALI Model. The questionnaire consisted of two parts. The first part was about personal information consisting of three questions: gender of the respondent, respondent's GPA, and the respondent's latest grade in the English subject. The other part which was concerning information related to students' opinions towards the instruction consisted of 12 Likert scale questions and an open-ended question. The process of constructing the questionnaire was guided by a book entitled '*Questionnaires in Second Language Research*' (Dörnyei, 2003).

The questionnaires were validated by three experts in the areas of EFL teaching and language assessment. A checklist (see Appendix L) was used for validating the questionnaire. There were two parts of this evaluation form. First, five 4-rating-scale items were employed for the quantitative data; and at the end of this part, the experts were asked to give their additional comments on their overall assessment of the questionnaire. The other part was composed of twelve semi-structured questions asking for experts' opinions on each item of the questionnaire; and at the end of this part, an open-ended question was used to ask for the experts' additional comments and suggestions. The experts' validation on the first part is quantitatively presented in the following table.

Table 3.9: Percentage of experts' opinions on the questionnaire

Items	Opinions				$\bar{X}$
	Excellent 4	Good 3	Moderate 2	Poor 1	
1. Instrument's explanation is clear.	2 (66.67%)	1 (33.33%)	-	-	3.67
2. Instruction is clear.	3 (100%)	-	-	-	4
3. Questions are purposeful and support the study.	3 (100%)	-	-	-	4
4. The language used is appropriate to The participants.	2 (66.67%)	1 (33.33%)	-	-	3.67
5. The format of the questionnaire is appropriate to the participants.	3 (100%)	-	-	-	4

Note: N = 3



All experts were satisfied with the questionnaire. There were a few comments and suggestions for editing given by the experts. For example, one expert suggested that question 12 should be separated into two questions. Moreover, another expert suggested that there should be some negative questions in order to confirm the reliability of the questionnaire. The questionnaires were revised accordingly. Finally, the revised questionnaire (Appendix K) was composed of 26 Likert scale questions and an open-ended question.

Piloting the instruments is essential. The appropriate time for a formal trial run is when a complete, almost final version of the questionnaire has been prepared (Dörnyei, 2003).

Isaac and Michael (1981) emphasizes the importance of pretesting a questionnaire. During the pretest, they advised that the following steps be taken.

- The sample group must be representative of the population for which the questionnaire is intended.
- Some space must be provided on the trial questionnaire for the respondents to make reactions and suggested changes.
- The conditions must be comparable to those expected in the main study.

The questionnaires were utilized in the pilot study with 48 Grade-10 students who participated in the ALI-Model lesson-plan verification. The items in the form of a 5-point rating scale were calculated for reliability values by using the method of the Cronbach's Alpha Reliability Estimate at the set point of 0.70. The Cronbach's Alpha Reliability obtained from the calculation was 0.794. No items were modified or deleted though the change or deletion of some items might have slightly increased the reliability value of the questionnaire. Such change or deletion was regarded as not being worthwhile because the possible increase in Alpha value would be slightly higher, whilst data obtained from such items were still considered as being useful.

#### 1.3.2.4 Interviews

Twelve randomly selected subjects from the experimental group in the main study were asked to participate in the interviews for more in-depth analysis and for confirmation of their opinions toward the instruction based on the ALI Model for enhancing secondary school students' English communicative abilities in large classes.

The semi-structured interviews were carried out with audiotape recordings on a one-to-one discussion basis. The interviews were conducted in Thai in order to obtain as much information as possible. The interviews were conducted within one week after the posttest. A set of questions, which was in accordance with the Questionnaire for Eliciting Students' Opinions, was designed as an interview framework and these questions asked students to give more details on what they thought about the instruction and also asked them to give some strengths and weaknesses of the instruction. Finally, they were asked to give comments and suggestions about how to improve the model.

### 1.3.2.5 Students' Reflective Journals

By the end of each unit, students would be asked to write a reflective journal showing their opinions toward what they had learned in the class. At first the students' reflective journal was designed as an open-ended question. That meant the students were allowed to write anything that they wished. However, after consulting the experts, they shared one comment in common which was that the researcher should provide some guided topics in the journal for students to follow. Accordingly, the revision of the students' reflective journal was made as well as the modified journal (see Appendix M) asking the students to show their opinions on what they had learnt in each step of the ALI Model. The students' reflective journal was presented in Thai in order to easily encourage students to do the journal.

All of the instruments for collecting the data were summarized and presented in the following table.

Table 3.10: The summary of the research instruments for collecting the data

Instruments	Objectives	Types of instruments	Time of distribution	Statistics
<b>1. The SLEP test (Forms 4 and 5)</b>	<ol style="list-style-type: none"> <li>1. To confirm that two groups of the subjects are comparable</li> <li>2. To assess students' English language proficiency</li> <li>3. To compare students' English language proficiency before and after the treatment</li> </ol>	A standardized test for use with students entering grades 7 through 12 or community colleges whose native language is one other than English	Before and after the treatment	<ol style="list-style-type: none"> <li>1. Descriptive Statistics</li> <li>2. t-test</li> </ol>
<b>2. Performance test</b>	To assess students' English communicative abilities	An alternative test developed based on the ALI Model to evaluate students' English communicative performance	After the treatment	<ol style="list-style-type: none"> <li>1. Descriptive Statistics</li> <li>2. t-test</li> </ol>

Instruments	Objectives	Types of instruments	Time of distribution	Statistics
<b>3. Questionnaire for eliciting students' opinions toward the ALI Model</b>	To collect students' opinions, comments, and suggestions to the instruction based on the ALI Model	1. Likert scale 2. Semi-structured	After the treatment	1. Mean score & S.D. 2. Content analysis
<b>4. Interview</b>	To collect student's in-depth opinions and comments on the ALI Model instruction	Semi-structured	After the treatment	1. Content analysis
<b>5. Students' reflective journal</b>	To collect student's general opinions, impressions, and comments on the ALI Model instruction	Guided topics	By the end of each unit	1. Mean score & S.D. 2. Content analysis

## 2. Experimentation of the ALI Model

The experiment was carried out with two comparable sample groups, which were randomly selected as an experimental group and a control group, at a secondary school in Nakhon Pathom Province in Semester two of academic year 2005. The experimental group's schedule was on Tuesday at 9:50 - 11:30 am while the control group's schedule was on Wednesday at 9:50-11:30 am. Both groups received the same content of a 10-unit communicative English course from the same instructor in the same classroom setting, but with different instructional models.

In the first week, the proficiency test (SLEP test form 4), as a pretest, was administered to the students of both the experimental and control groups and their scores were collected. Later, both classes received the treatment for 10 weeks, one using the ALI Model instruction, and the other using the PPP Method instruction. By the end of each lesson in the experimental group, the students were asked to write a reflective journal in Thai to express their opinions toward the lesson, a so-called one-minute paper (Angelo & Cross, 1993: 148), which took the last five minutes of the class for students to say their opinions about the instruction. Sutherland (1996: 86) suggests that it is important to give students frequent opportunities to provide feedback about the class when using new approaches. After week eight, the performance test (reading and writing tasks) was assigned to the students of both groups in order to pave the way for the presentation task by the end of the course. The students' writing assignments and presentations were graded according to the rubrics by two teachers in order to confirm the reliability of the scores. The students' scores were then collected.

After ten weeks of the treatment, the presentation of the performance test was done in both groups as the listening-speaking performance tests. The scores from their performance were graded according to the rubric by two teachers in order to confirm the reliability of the scores. The scores were then collected. After the performance test, the questionnaire for eliciting students' opinions towards the ALI model instruction employed all throughout the course was distributed to the students in the experimental group in order to elicit their opinions towards model. The responses from the students were collected. Moreover, twelve students were randomly selected to be interviewed and asked for opinions regarding the instruction. Finally, a parallel form of the proficiency test (SLEP test form 5) was administered to the students of both groups as the posttest. The students' posttest scores were then collected.

### 3. Analysis of the effectiveness of the ALI Model

#### 3.1 Data analysis for research questions 1 – 2

The first two research questions were concerned with the effects of the ALI Model instruction on students' English communicative proficiency scores. The independent variable (IV) was the two modes of instruction: the ALI Model instruction used in the experimental group, and the PPP Method instruction used in the control group. The dependent variable (DV) was the scores on the proficiency and performance tests. To analyze the data, the t-test (Hatch & Farhady, 1982), using the independent-samples test, was conducted to determine the differences between the posttest scores of the proficiency test between groups. In addition, the independent-samples test was used to analyze the performance test scores between the control and experimental groups.

#### 3.2 Data analysis for research questions 3

Research question three was regarding students' opinions on the ALI Model instruction implemented in large classes for enhancing students' communicative abilities. The data from the questionnaire for eliciting students' opinions toward the ALI Model was analyzed using  $\bar{x}$  and S.D. The data from students' reflective journals was analyzed qualitatively by content analysis in order to collect students' opinions, suggestions, and comments, and also analyzed using percentage and  $\bar{x}$  to obtain the students' opinions on the steps of the ALI Model. The

data from the interviews were analyzed by content analysis and percentage to get the opinions and suggestions toward the instruction.

#### 4. Conclusion

After the data analysis, the findings of the experiment were concluded. The result from the study was used to revise the ALI Model, show the effectiveness of the model, and express the students' opinions toward the ALI Model instruction.

### **Summary**

This study is a developmental and experimental research. The research has been conducted in two main phases as follows:

Phase 1: The development of the Active Learning Instructional (ALI) Model for enhancing secondary school students' English communicative abilities in large classes

Phase 2: The evaluation of the effectiveness of the Active Learning Instructional (ALI) Model for enhancing secondary school students' English communicative abilities in large classes.

The experiment was based on quasi-experimental research design. The study was conducted with two groups of 96 Grade-10 students for 14 weeks. The experimental group obtained a 10-week English communicative course based on the ALI Model instruction while the control group received a 10-week English communicative course based on the PPP Method instruction. After the experiment, the English proficiency scores between groups were compared. Besides this, the English performance scores of the two groups were compared. In addition, three main types of the instruments—questionnaire, interview and students' reflective journals—were utilized to collect the data of the students' opinions on the ALI Model instruction using  $\bar{X}$ , S.D., frequency, percentage and content analysis.