

Chapter 3

Research Methodology

3.1 Introduction

The inclusion of language assessments in second language learning contexts may improve second language acquisition for some learners and not others. This study measured the effects of a particular assessment strategy on the ability and opinions formed by upper secondary learners of English in Thailand relative to age, pre-instructional proficiency and gender. The procedures carried out are discussed here.

3.2 Population and Sample

Participants were native speakers of Thai learning English as a foreign language in Thailand at an upper secondary national school in rural Northern Thailand in secondary levels three through six; ages ranged from 13 to 19 ($\mu=15.73$). 416 students from twelve class sections participated in the study, 318 in nine sections of whom participated in the main study, 147 in the treatment and 171 in the control. The remaining participants ($n=98$) were reserved to pilot instruments.

Table 5 Participants

	Level	Section	Average Age	Age Classification	Class Size	Assignment
1	3	1	14.73	Younger	34	Treatment
2	3	2	14.39	Younger	30	Pilot-Control
3	3	3	14.5	Younger	36	Pilot-Treatment
4	3	4	14.55	Younger	31	Control
5	3	5	14.44	Younger	35	Treatment
6	3	6	14.67	Younger	34	Control
7	4	1	15.63	Younger	32	Pilot-Treatment
8	4	4	15.54	Younger	33	Control
9	5	1	16.38	Older	38	Treatment
10	5	5	16.42	Older	34	Control
11	6	1	17.56	Older	39	Control
12	6	2	17.3	Older	40	Treatment
Total Treatment			15.71	Younger	147	Treatment

	Level	Section	Average Age	Age Classification	Class Size	Assignment
Total Control			15.75	Younger	201	Control
Total			15.51	Younger	416	12
Total Pilot Participants			14.84	Younger	98	3
Total Active Participants			15.73	Younger	318	9
Average			15.51	Younger	35	5

Prior to instruction, all learners received a pre-instructional placement examination of ability and results were used to sort learners into four bands of equal proficiency from which equal numbers of learners ($n=25$) were randomly selected from treated and control groups. Table 6 represents the percentile divisions relative to initial scores on the placement examination used to sort learners into four equal bands of proficiency.

Table 6 Range Percentiles

N	Valid	200
	Missing	0
Percentiles	25	31.25
	50	36.25
	75	39.75

Table 7 represents the proficiency ranges and numbers of participants selected for each proficiency range relative to treatment group, reflecting in sum the composition of the data sample

Table 7 Sample

Data Based on Performance Examination Scores							
Proficiency Range	Range		Range		Number of Students per Band		Total No. of Students
	(Mean)		(%)				
Description	Hi	Lo	Hi	Lo	Treatment	Control	Total
Lowest	21	31.25	46.67	69.44	25	25	50
Lower-Middle	31.26	36.25	69.47	80.56	25	25	50
Upper-Middle	36.26	39.75	80.58	88.33	25	25	50
Highest	39.76	45	88.36	100	25	25	50
Total	21	45	46.67	100	100	100	200

Table 8 depicts similar information with populations by gender also represented.

Table 8 Purposeful Sample by Gender

Proficiency Range	Treatment				Control				Total
	#	Male	Female	Unknown	#	Male	Female	Unknown	#
Lowest	25	13	11	1	25	7	12	6	50
Lower-Middle	25	8	17	0	25	4	18	3	50
Upper-Middle	25	6	19	0	25	6	18	1	50
Highest	25	7	18	0	25	5	17	3	50
Total	100	34	65	1	100	22	65	13	200

3.3 Treatment Instruments

Participants were randomly assigned by section to receive treatment or not. Sections selected to receive treatment were issued a series of short, predominantly multiple-choice assessments at the end of every lesson. Learners assigned to the control group received no assessment.

3.3.1 Teaching Methods

Instruction took place in three weekly 50-minute sessions for a total of 150 minutes. All learners received the same instruction, which endeavored to address certain microskills useful to the art of conversational English. In the first class, learners watched and discussed an American film, focusing both on comprehension of the narrative and its dialogue and how to discuss them coherently. In the second class, learners practiced providing and understanding descriptions of people based upon vocabulary items presented at the beginning of the lesson. In the third class, learners listened to and discussed poetry, finally focusing on a single sonnet by Robert Frost. Table 9 gives an overview of the content of instruction.

The instructor built classes around fixed-length learning materials, such as the film in the first lesson, the presentation in the second lesson, and the poems in the third lesson, in order to improve the replicability of lessons and therefore the reliability and similarity of class instruction. An American film, *Twilight* (Hardwicke, 2008), based upon the first in a series of young-adult novels (Meyer, 2005), was chosen primarily for its demonstrated appropriateness for the age of participants. The film combines the genres of romance and horror, both of which are independently popular in Thailand and also popular when combined (Sukwong, 2001). The characters in the film are of similar ages as the students in the class (upper secondary levels). The language used in the film is sparse and clearly enunciated as well as fairly undemanding of beginning learners. A film, in particular, was used to bridge the gap in schema potentially lacking in learners whose own knowledge backgrounds were likely somewhat different from the instructor and course designer. Poems, similarly, were

chosen as a fruitful basis for discourse analysis and to provide a mutual basis for discussion (Thornbury, 2005).

Table 9 Instruction

Class #	Instructional Materials	Language Skills		Activity	Microskill
		Skill	Application of Skill		
1	An American film, <i>Twilight</i> .	Listening	Listening to dialogue in film for comprehension of narrative and dialogue.	Watching a Film	Listening to dialogue and comprehending narrative structure
		Speaking	Speaking selected excerpts from film dialogue.		
		Discussion	Discussing film events, characters, plot and aesthetics.		
2	A slideshow of vocabulary (descriptive characteristics) and corresponding illustrations	Listening	Listening to descriptions of people and vocabulary items.	Providing and Understanding Descriptions	Listening for cues related to appearance and recognizing people based upon cues; providing comprehensible descriptions of people
		Speaking	Practicing using vocabulary.		
		Performance	Providing oral descriptions of people.		
3	<i>Acquainted with the Night</i> , a sonnet by Robert Frost	Listening	Listening to different selections of poetry and noticing the language and rhythms.	Participating in Oral Recitations of Poetry	Identifying words in poems by sound and context
		Speaking	Repeating the language used in poems.		
		Discussion	Discussing the meanings and language used in poems.		

The total time of instruction was 150 minutes. Table 10 describes allocation of time.

Table 10 Allocation of Instructional Minutes

#	Instructional Materials	Language Skills	Activities and Micro-skills	Treatment				Control			
				<i>Ad*</i>	<i>Ma*</i>	<i>Di*</i>	<i>As*</i>	<i>Ad*</i>	<i>Ma*</i>	<i>Di*</i>	<i>As*</i>
				<i>Time in Minutes</i>				<i>Time in Minutes</i>			
1	An American film, <i>Twilight</i> .	Listening	Watching a film; understanding and discussing dialogue and narrative.	10	25	5	10	0	40	10	0
		Speaking									
		Discussion									
2	A slideshow of vocabulary and illustrations	Speaking	Learning vocabulary; providing and understanding descriptions of people.	10	20	10	10	0	30	20	0
		Listening									
		Discussion									
3	<i>Acquainted with the Night</i> , a sonnet by Robert Frost	Listening	Listening to and discussing poetry; identifying individual words within poems.	10	20	10	10	0	30	20	0
		Speaking									
		Discussion									
Average				10	21.67	8.33	10	0	33.33	16.67	0
Total				30	65	25	30	0	100	50	0
%				20	43.33	16.67	20	0	66.67	33.33	0
4	Placement Examination			10	0	0	40	10	0	0	40
5	Performance Examination			10	0	0	40	10	0	0	40
Average				10	13	5	22	4	20	10	16
Total				50	65	25	110	20	100	50	80
%				25	32.5	12.5	55	10	50	25	40
<i>Ad*</i>		<i>Ma*</i>		<i>Di*</i>		<i>As*</i>					
Administration		Materials		Discussion		Assessment					

3.3.2 The Function of Treatment Instruments

The treatment instruments were designed to measure the skills and micro-skills introduced and practiced within the same lesson. The nature of the treatment instruments were as follows:

- 1 Comprehension of Media: The first treatment instrument was used as a means of reviewing the material discussed in class in which a film was shown and discussed concurrent to viewing. The instrument measured retention of items discussed during the class. In control classes, the information was only discussed and not tested.

- 2 Providing Descriptions: The lesson introduced learners to vocabulary necessary for providing descriptions of human beings, including facial and bodily characteristics and general qualities. The treatment instrument tested learner ability to recognize the vocabulary practiced in the lesson and to understand descriptions and provide descriptions in written form.
- 3 Recognizing Words in Poetry: In this lesson learners listened to and discussed several poems briefly before focusing on discussion of a sonnet by Robert Frost in detail. The treatment instrument consisted of a multiple-choice format cloze passage of the sonnet for which learners were to select missing words from a list of choices given, a task that would be repeated on the final performance examination with a different poem.

3.3.3 The Design of Treatment Instruments

The treatment instruments were designed to contain a fixed number of items each, with the second treatment instrument containing the greatest number of items and highest degree of difficulty. The structure of the treatment instruments was as follows:

Table 11 Treatment Instruments

#	Instrument	Language Skill	Ability-Task	Micro-skill or Content	Formats	#	Value
1	Film Analysis	Listening	Watching and discussing a film.	Comprehending narrative and reiterating topical knowledge.	Multiple-choice	5	5
2	Appearances	Speaking and Listening	Providing and understanding descriptions of people.	Providing and understanding descriptions of people.	Multiple-choice and short-answer	10	10
3	A Poem	Listening	Listening to and discussing poetry.	Listening to identify individual words in context.	Multiple-choice cloze test	5	5
<i>Total</i>						20	20

Treatment instruments used multiple-choice format with some short-answer questions on the second treatment instrument and consumed the final ten minutes of the sessions in which they were administered. Treatment instruments were graded and returned to learners at the beginning of class the following week.

3.3.4 The Validity and Reliability of Treatment Instruments

Two experts in the field of TESOL reviewed treatment instruments prior to administration and changes were made based upon their suggestions. The data resulting from participants of the study was analyzed for reliability and found to possess good internal consistency. Table 12 presents a summary of reliability analyses for treatment instruments.

Table 12 Treatment Instrument Reliability

Measure	Treatment Instrument #1	Treatment Instrument #2	Treatment Instrument #3
Cronbach's Alpha*	0.72	0.77	0.74
Split-Half (odd-even) Correlation	0.75	0.85	0.91
Spearman-Brown Prophecy	0.86	0.92	0.96
Mean for Test	3.54	8.04	4.24
Standard Deviation for Test	1.29	1.73	1.16
* $0.7 \leq \alpha < 0.9$	Good (Low-Stakes testing)		

3.3.5 Consequentiality of Treatment Instruments

Treatment instruments represented twenty points and 5% of learners' total grades in the class, qualifying for this and other reasons, i.e., Thailand's no-fail policy (Halligan, 2011), as low-stakes assessments. Total weighted scores are presented in Table 13.

Table 13 Value of Instruments

Measure	Grade %
Defined by Local Context	50%
Final Examination	25%
Midterm Examination	20%
Treatment Instruments/Participation	5%
<i>Total</i>	<i>100%</i>

3.4 Data Collection and Instruments

This section describes the instruments used to collect data, of which there were five:

Table 14 Data Collection Instruments

	Instrument	Quantitative	Qualitative
1	The English Language Placement Examination	×	
2	The English Language Performance Examination	×	
3	Treatment Instruments	×	
4	The Opinion Survey Questionnaire	×	×
5	The Field Journal	×	×

3.4.1 The English Language Placement Examination

The English Language Placement Examination was issued prior to instruction and designed to measure general knowledge of English relative to topical content (e.g., prepositions and colors) and certain and general ability tasks (e.g., understanding the gist of a conversation). Results of the examination were used to place learners within four ranges of proficiency used to compare learner performance on the final examination.

3.4.1.1 Description of The English Placement Examination by Section

The placement examination consisted of seven sections, five multiple-choice, one dictation section and one written response section. The examination contained a maximum value of forty-five points. The examination was issued in one 50-minute period. All questions and answer stems for multiple-choice questions were read aloud in English. The nature of the class was conversational English and could not presuppose literacy. Consequently, questions were read out loud.

Table 15 Overview of Placement Examination Sections

	<i>Title</i>	<i>Skill</i>	<i>Micro-skill or Content</i>	<i>Format</i>	<i>Items</i>	<i>Value</i>
1	Prepositions and Directions	Vocabulary	Prepositions and Directions	Multiple-choice	10	10
2	Colors	Listening Vocabulary	Colors	Multiple-choice	10	10
3	Seasons in America	Sociolinguistic	Seasons	Multiple-choice	5	5
4	At the Café	Listening	Content	Multiple-choice	5	5
5	Listening	Listening	Dictation	Writing	5	5
6	At the Hospital	Listening	Gist	Multiple-choice	5	5
7	Writing	Writing	Expressing Aims	Writing	5	5
	<i>Total</i>				45	45

The First Section of The Placement Examination

On the first section of the placement examination learners were shown images on slides of locational and directional relationships and asked to choose the correct preposition or direction provided. All questions and answers were read out loud in English.

The Second Section of The Placement Examination

On the second section of the placement examination, unambiguous colors were projected onto the projection screen and students identified each color from a selection of colors provided. All questions and answers were read out loud in English.

The Third Section of The Placement Examination

In the third section of the placement examination, students identified photographs with associated names of seasons in North America to measure topical knowledge of seasons in North America. All questions and answers were read out loud in English.

The Fourth Section of The Placement Examination

On the fourth section of the placement examination, students listened to a recorded dialogue several times and selected answers to questions regarding content in the dialogue. All questions and answers were read out loud in English.

The Fifth Section of The Placement Examination

On the fifth section of the placement examination, students transcribed a single sentence read aloud several times. All questions and answers were read out loud in English.

The Sixth Section of The Placement Examination

In the sixth section of the placement examination, students listened to a recorded dialogue and selected the omitted words from a list provided (cloze). All questions and answers were read out loud in English.

The Seventh Section of The Placement Examination

On the seventh section of the placement examination, students provided written responses to a question posed in English. The question was read and aloud and projected as text. Information on grading occurs in the appendices. All questions and answers were read out loud in English.

3.4.1.2 Validity and Reliability of The English Placement Examination

Two professionally qualified experts in the field of TESOL reviewed the examination to ensure its validity and reliability. The examination was piloted to 91 learners whose data would not be included in the final study and whose personal characteristics were similar to the final sample.

Table 16 Placement Examination Normality and Reliability

	N of Items	N of Students	Average Age	Anderson-Darling	Cronbach Alpha α
Pilot	45	91	15.13	(p)	0.96**
Actual	45	200	15.77	0.1	0.92**
** $\alpha \geq 0.9$				Excellent (High-Stakes testing)	

3.4.2 The English Language Performance Examination

The English Language Performance Examination measured final language ability in the class relative to microskills introduced and practiced during the course of learning.

3.4.2.1 Description of the Performance Examination by Section

The final performance examination contained five sections; each section was designed to measure a different task-dependent ability. Examination tasks referenced lesson material, except for the third section, which measured learner performance on an unfamiliar task. All instructions and answers were read out loud in English. An overview of tasks and formats is presented in the table below (Table 17).

Table 17 Final Performance Examination Sections

	<i>Section Title</i>	<i>Language Skills</i>	<i>Micro-Skills or Content</i>	<i>Format</i>	<i>Items</i>	<i>Value</i>
1	<i>Twilight</i>	Topical Knowledge + Listening	Retention of information related to content and class discussion	Multiple-choice	6	6
2	Listen to the Passage	Listening	Listening for Words	Multiple-choice	6	6
3	What is this Newscast About?	Listening	Watching and listening for the gist of a newscast	Multiple-choice	6	6
4	Describe Your Friend	An Approximation of Speaking	Providing descriptions of people	Written Responses	6	6
5	Draw What You Hear	Listening	Understanding descriptions of people	Picture	6	6
	<i>Total</i>				30	30

The First Section of the Performance Examination

The first section of the final performance examination measured learner apprehension of the lesson content presented in the first lesson of the instructional phase in a multiple-choice format. Section one contained one listening task for which learners were instructed to identify a character from the film by the sound of his voice alone. All questions and answers were read out loud in English.

The Second Section of the Performance Examination

The second examination task measured ability to discern words contained in poems. A recording of a poem was played and the poem itself was printed on a page for which words had been deleted. This section was read out loud in English; the poem was read out loud five times and the answer keys were read out loud twice.

The Third Section of the Performance Examination

The third examination task measured comprehension of a newscast, its gist and details contained within the narrative. A similar task was not included in the course of instruction. All questions and answers were read out loud in English.

The Fourth Section of the Performance Examination

The fourth examination task measured ability to describe humans: students composed sentences to describe a photograph of person using words provided on the examination paper. All questions and answers were read out loud in English.

The Fifth Section of the Performance Examination

The fifth examination task measured ability to understand descriptions: learners drew pictures of a person based upon what they heard, emphasizing each quality in their illustration. All questions and answers were read out loud in English.

3.4.2.2 Validity and Reliability of The English Language Performance Examination

Two experts in the field of TESOL reviewed the performance examination for reliability and validity. It was piloted to 61 students and was shown to have high internal consistency. Reliability is depicted in Table 18.

Table 18 Final Performance Examination Normality and Reliability

	N of Items	N of Students	Average Age	Anderson-Darling	Cronbach Alpha α
Pilot	30	61	15.13	(p)	0.94**
Actual	30	200	15.77	0.117	0.88*
$**\alpha \geq 0.9$				Excellent (High-Stakes testing)	
$*0.7 \leq \alpha < 0.9$				Good (Low-Stakes testing)	

3.4.3 The Opinion Survey Questionnaire

The Opinion Survey Questionnaire measured the opinions learners formed about the language course using a twenty-five statement, five-point Likert-scale of agreement. Interpretations of the rating scale are presented in Table 19 and Table 20.

Table 19 Opinion Survey Questionnaire Rating Scale

<i>Rating</i>	<i>Interpretation</i>
1	Strongly Disagree
2	Disagree
3	Neither Agree Nor Disagree
4	Agree
5	Agree Strongly

Table 20 Opinion Survey Questionnaire Rating Interpretation

<i>Range</i>	<i>Interpretation</i>
4.21–5.00	Strongly Agree
3.41–4.20	Agree
2.61–3.40	Neutral
1.81–2.60	Disagree
1.00–1.80	Strongly Disagree

3.4.3.1 Description of The Opinion Survey Questionnaire by Section

The first section measured general opinions toward language learning and the class. The second section measured opinions regarding lessons. The third section measured opinions regarding the assessment plan. The fourth section measured opinions toward the teacher. The fifth section measured learner confidence.

Table 21 Opinion Survey Questionnaire Sections

	<i>Measure</i>	<i>Responses</i>	<i>Items</i>	<i>Value</i>
1	Opinion of class in general	1–5	5	25
2	Opinion of lessons	6–9	4	20
3	Opinion of assessment	10–14	5	25
4	Opinion of teacher	15–18	4	20
5	Confidence	19–20	2	10
	<i>Total</i>		20	100

3.4.3.2 Validity and Reliability of The Opinion Survey Questionnaire

Two experts in the field of TESOL examined the survey prior to issuance for consistency and it was piloted to 55 learners and shown to have good internal consistency.

Table 22 Opinion Survey Questionnaire Internal Consistency

	N of Items	N of Students	Average Age	Anderson-Darling	Cronbach Alpha α
Pilot	20	55	15.13	(<i>p</i>)	0.92*
Actual	20	220	15.75	0.317	0.91**
** $\alpha \geq 0.9$				Excellent (High-Stakes testing)	

3.4.4 Field Journal

A field journal was used to record instructor observations in qualitative and quantitative modes. In addition to incidental, unplanned notations, review notations were made retrospective of every lesson. The field journal also contained a quantitative element that measured teacher perception of student investment in the class relative to the following five statements and a four point Likert-scale:

Table 23 Field Journal Observation Checklist

		Not at All	Somewhat	Adequately	Well
Statement		0	1	2	3
1	Students are engaged in every task.				
2	Students completed every task assigned.				
3	Students appear to be challenged.				
4	Students appear to be enjoying themselves.				
5	Students are collaborating.				

3.5 Data Collection

Prior to instruction, class sections were randomly sorted into sections that would receive treatment (n=5) and sections that would not receive treatment (n=5) using simple random, roll-of-the-dice sampling. In the first session, all learners were issued the placement examination. Instruction began on session two and continued through session four. During sessions two through four, learners who were assigned to the treatment group were issued treatment instruments at the end of each session. On the fifth and final session, the final performance examination and opinion survey were administered. Field notes and observations were made throughout instructional periods as well as directly afterwards. More information on the exact sequence of data collection procedures can be viewed in Table 24.

Table 24 Summary of Data Collection Procedures

	Instrument	Time of Collection			Type of Data	
		1 Prior to Instruction	2 During Instruction	3 Following Instruction	Quant.	Qual.
1	The English Language Placement Examination			x	x	

	Instrument	Time of Collection			Type of Data	
		1 Prior to Instruction	2 During Instruction	3 Following Instruction	Quant.	Qual.
2	The English Language Performance Examination	x			x	
3	Treatment Instruments (3)		x		x	
4	The Opinion Survey Questionnaire			x	x	x
5	The Field Journal		x		x	x

3.6 Data Analysis

Independent samples *t*-tests in SPSS 17 (two-tailed) were used to determine the whether differences in means of scores were significant between treated and untreated learners and by age, proficiency and gender. Comparisons of means were conducted relative to the scores of participants on the final examination, treatment instruments, the opinion survey and also certain sections of the placement examination. Reliability analyses using Cronbach's alpha were conducted in SPSS 17. Correlations between instruments were measured using Pearson's correlation coefficient calculator in SPSS 17. Quartiles used for sorting learners by proficiency were generated using SPSS 17 and Excel 2011. Anderson-Darling tests in XLSTAT-Pro software (Addinsoft, 2014) were used to verify the normality of distributions. A reliability calculator Excel spreadsheet generated by Del Siegle (Siegle, 2014) was used to generate some reliability statistics charts. Q-Q Plots, boxplots and histograms were produced in SPSS 17. GraphPad QuickCalcs were also used for additional shorthand calculations (GraphPad Software, 2014). Table 25 summarizes data analysis, instruments and programs.

Table 25 Data Analysis

Procedure	Instruments	Factors	Coefficient	Statistical Process	Program
Comparing Means	Performance Examination	Treatment	Probability Value (<i>p</i>)	Independent samples <i>t</i> -tests (2-tailed)	SPSS
	Treatment Examinations	Age			

Procedure	Instruments	Factors	Coefficient	Statistical Process	Program
	Opinion Survey	Proficiency			
	Placement Examination	Task			
		Gender			
Correlation	Performance Examination		Pearson's product moment correlation coefficient (r)	Pearson Correlation Measurement	SPSS
	Treatment Examinations				
	Placement Examination				
Reliability	Performance Examination		Cronbach's Alpha (α)	Cronbach's Alpha Reliability Assessment	SPSS, Excel add-on
	Treatment Examinations				
	Opinion Survey				
	Placement Examination				
Normality	Performance Examination		Normality Alpha (α)	Anderson-Darling test for Normality	XLStat Pro
	Treatment Examinations				
	Opinion Survey				
	Placement Examination				