

CHAPTER THREE

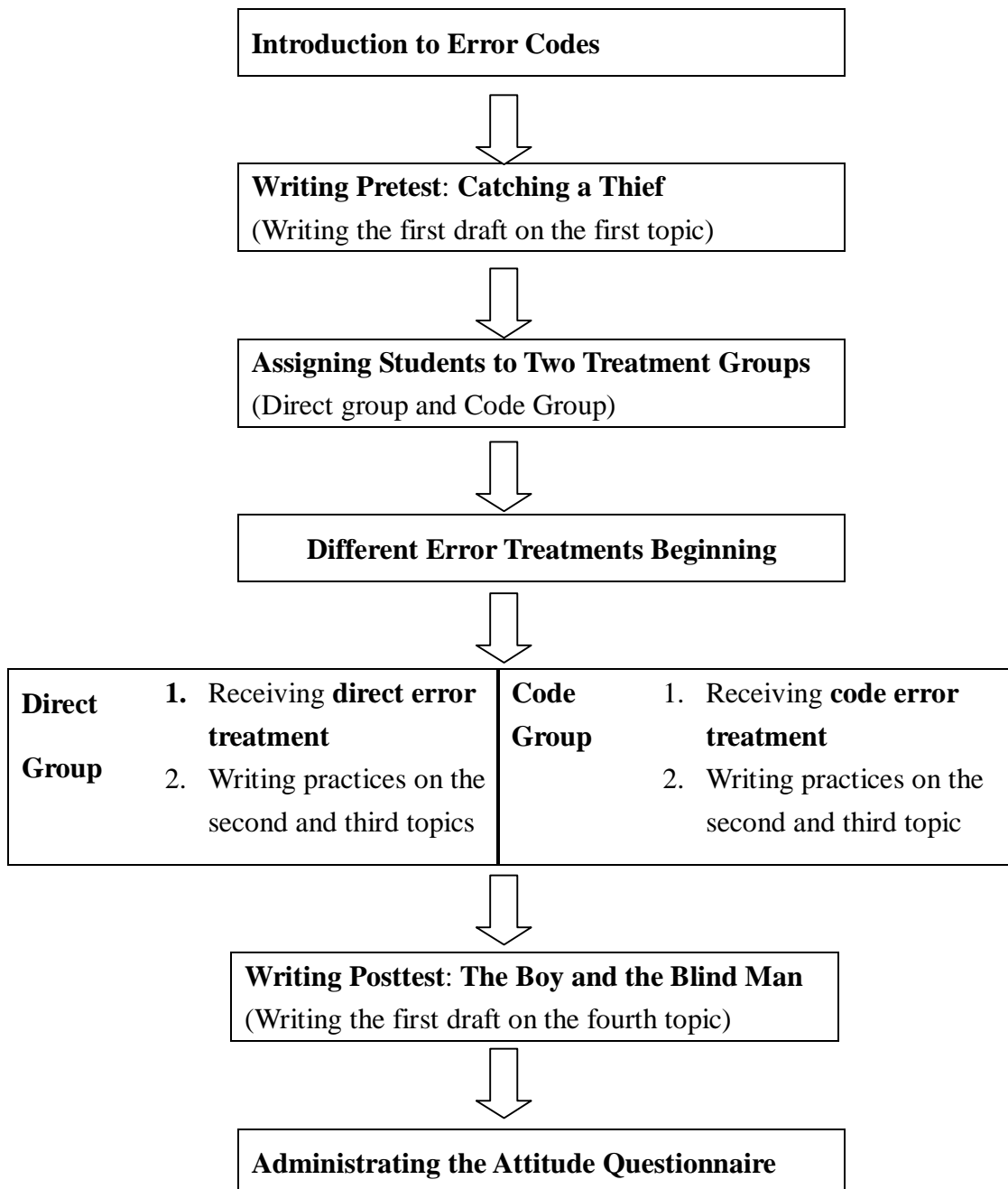
METHOD

In this chapter, the method employed to conduct the present study will be described. It consists of four sections. In the first section is the description of the research design. In the second section the procedures of the experiment are reported. In the last two sections, the instruments employed for the study as well as the procedures used to collect and analyze the data are discussed.

Research Design

This empirical study was designed to investigate the differences between the effects of two error correction methods on the writing quality and accuracy of students at two levels of writing proficiency. It also explored students' attitudes toward the treatments they received. A total of 90 students who fell into two writing proficiency levels were recruited and assigned to two treatment groups, each of which contained 45 students. To demonstrate the procedures of the experiment, a brief flow chart was displayed in Figure 1.

Figure 1
The Flow Chart of the Procedures of the Experiment



In this experiment, all students in two treatment groups were required to write eight drafts on four topics, for each of which two drafts were produced (see Appendix A for the four topics). The first draft on the first topic was the pretest writing administered at the beginning of the experiment, while the first draft on the fourth topic served as the writing posttest of this study.

After the pretest, the different treatments began. For each draft written, the code correction group received feedback in which the types of errors were identified with a coding system designed by the researcher (See Appendix B for the coding system), whereas the direct correction group received feedback with correct answers provided for their errors. At the end of the experiment, the attitude questionnaire was administered to probe into students' perceptions of the treatment they received and their attitudes toward it.

Setting

The current study was conducted in National Chiayi Girls' Senior High School in Chiayi City and the participants in the study were the researcher's students. The research site was chosen mainly because recruiting her own students enabled the researcher to understand firsthand how the students were learning (Perl, 1980; Zamel, 1983) and to monitor their progress throughout the experiment. Another benefit was that the researcher was able to build mutual trust and better relationship with the participants, which were needed for the smooth progress of the experiment.

Participants

The participants in this study were 90 female seniors from two classes in National Chiayi Girls' Senior High School. Forty-three of them were from a science class while the other forty-seven students belonged to a social science class. The researcher decided to recruit the senior students because they were equipped with better grammatical knowledge and they were more proficient in writing than the junior and sophomore

students.

During the period of time when the experiment was being conducted, the science class received 6 hours of English instruction every week. On average, 4.5 hours were spent on the teaching of the textbook (Far East New English Reader for Senior High School, Book Five) and the remaining 1.5 hours were mainly spent on the practice of writing and the training of the writing skills. The students in the science class had 7 hours of English instruction in total. Apart from the 6 hours of the same instruction as that of the science class, an extra hour of English class was spent on reviewing the materials they had been taught previously. Most of the time, the review took the form of a simulated test, with the teacher's clarification of students' erroneous answers at the end of the test.

Since the teaching hours of two classes were different, before the start of different treatments, the researcher mixed the students from two classes to eliminate the influence of the unequal teaching hours. She first assigned 90 students to higher- and lower-proficiency groups, making each group contain 45 students. The placement was based on the pretest writing scores graded by two independent raters ($r = .86$, $p < .05$). Then, the students in each proficiency group were randomly assigned to two treatment groups of 45 students. The students in the experimental group received code error correction while those in the control group received direct error feedback. The results of the placement were displayed in Table 1.

Table 1
Distribution of the Participants

Treatment types proficiency	Direct Correction Group		Code Correction Group	
	Social Science	Science	Social Science	Science
Higher Proficiency	11	11	11	12
N = 45	22		23	
Lower Proficiency	13	10	12	10
N = 45	23		22	

Note: Social science class: $N = 11+13+11+12 = 47$; Science class: $N = 11 +10 +12 +10 =43$

Table 1 showed that the direct correction group was composed of a total of 45 students, 24 of which were from social science class and 21 were from science class. The code correction group contained also a total of 45 students, 23 of which belonged to social science class and 22 were from science class.

After the distribution, the students' pretest writing scores were subjected to a MANOVA procedure to test if there was any significant difference between the mean scores of two treatment groups. The result was displayed in Table 2.

Table 2
Comparison of Mean Scores of Two Groups in the Pretest

Source of Variation	SS	F	Sig of F
Treatments by MWITHIN 'Pretest Score'	.91	.11	.744

According to Table 2, no significant difference was found between the mean scores in the pretest for the two treatment groups at the .05 level ($F = .11, p = .744 > .05$). This result indicated that there was no significant difference between the

writing abilities of the two groups at the beginning of the experiment.

Procedures of the Experiment

The entire experiment was conducted from October 10, 2005 to January 29, 2006. Before the beginning of the experiment, all the students received ten class meetings of writing instruction, preparing them for the writing practice. After that, writing practice and different treatments began. At the end of the experiment, their attitudes toward the treatments were surveyed by a questionnaire (see Appendix C and D for the questionnaire).

Writing Instruction before the Experiment

The researcher's writing instruction emphasized both form and content. Since all the students had received grammar instruction in the first two academic years, the researcher only gave them a brief review of the grammar they had learned, and introduced to them the most important sentence patterns in English. Besides the grammar instruction, the researcher also taught students how to organize a composition. Before the writing practice, students were introduced to important concepts including pre-writing and outlining, structure of a paragraph, unity and coherence in writing, and the concept of narrative genre. The pre-experiment writing instruction equipped students with principles guiding them to compose a narrative. This instruction also provided students with the grammar knowledge and concepts they needed for decoding the symbols used in code error correction for categorizing the errors they made when the treatment began.

Writing Procedures and the Treatments

The experiment began on October 11, 2005, and lasted for 16 weeks. The writing procedures included, first, the teacher's instruction of error codes; second, the writing practice; third, the different error feedback approaches; and fourth, the completion of the error checklists. The procedures were first presented in detail in Table 3 and then

reported in the following sections.

Table 3

Procedures of the Experiment

Week	Procedures	Drafts & Rewrites
1 st - 2 nd	Introducing the list of error codes Explaining how to use the error checklist	Pre-treatment
3 rd	Writing the draft on the first topic (D1)—the writing pretest	Draft 1
4 th	Assigning 90 students to two treatment groups at two proficiency levels Beginning two different treatments	
5 th	Filling out the error checklist of D1	Rewrite 1
	Revising D1 to complete its rewrite (R1)	
6 th	Filling out the error checklist of R1	Draft 2
	Writing the draft on the second topic (D2)	
7 th	Filling out the error checklist of D2	Rewrite 2
8 th	Revising D2 to complete its rewrite (R2)	
9 th	Filling out the error checklist of R2	Draft 3
	Writing the draft on the third topic (D3)	
10 th	Filling out the error checklist of D3	Rewrite 3
11 th	Revising D3 to complete its rewrite (R3)	
12 th	Filling out the error checklist of R3	Draft 4
13 th	Writing the fourth draft on the fourth topic (D4)—the writing posttest	
14 th	Filling out the error checklist of D4	Rewrite 4
15 th	Revising D4 to complete its rewrite (R4)	
		Filling out the error checklist of R4
16 th	Administrating the attitude questionnaire	

Introduction to Error Codes

At the pre-treatment stage—the first two weeks of the experiment, the researcher introduced the coding system she designed to all students, explaining the meaning of every symbol by referring to the examples listed in the table (see Appendix B for the coding system). Then she distributed the error checklists to students, demonstrating how this checklist would be used in the future writing practice (see Appendix E and F for the checklist). She explained to them that when they received their composition graded by the teacher, they had to fill out their checklist according to the types of errors they made.

Writing Practice

Throughout the experiment, a total of 16 hours were spent on writing practice. Eight drafts in all were written on four topics based on the pictures given to them (See Appendix A for the topics and the pictures). For each topic, each student had to produce two drafts—the first and the second draft. Based on the teacher's feedback on their first draft, they were requested to compose their second draft.

Each of the four topics was about a narrative description of a particular event. Students were asked to tell the event or stories by describing the given pictures. The choice of picture descriptions, rather than providing a topic only, was to limit the variety of the content which was not measured in this study.

Different Error Correction Methods

After the pretest writing, the two treatment groups began to receive error feedback from their teacher. For both groups, all the errors appearing in the writing were corrected by the teacher. There were two reasons for the practice of comprehensive correction. First, the faulty linguistic structures, if not identified, may become ingrained in the student's interlanguage system (Lalande, 1982). Second, since students preferred comprehensive correction, writing teachers had to tailor to

their needs so students' motivation was not decreased (Ferris, 2004; Lee, 2004).

Although all students received comprehensive correction in their writing, students of two treatment groups received different error feedback methods differing in the degree of explicitness. On each draft written, the students of the direct correction group received the teacher's feedback in which the correct linguistic forms of grammatical errors were provided and the number of the errors was indicated by red tallies marked above the errors (see Appendix G for the sample composition of a student from the direct correction group). Students in the direct correction group had to rewrite their first draft on each of the four topics according to the correct answers teachers provided for them. Usually, they merely copied down the models for the erroneous usages, together with the rest correct sentences, to complete their second draft.

Compared to the direct correction group, the revision of the first draft demanded more efforts from the students in the code group. For the errors in their writing, they received feedback from their teacher in which only the types of error were identified according to the Table of Error Codes (see Appendix B for the Table of Error Codes) but no correct forms were offered to them. The number of errors was also indicated by red tallies marked above the errors (see Appendix G for students' sample compositions). When they rewrote their first drafts, they had to refer to their Table of Error Codes to decode the meaning of the codes with which their teacher addressed their errors. When they understood what the codes meant, they could refer to the grammar textbooks or the dictionary for the correct linguistic forms of the errors indicated. When the correct answers were found, they then incorporated them to complete their second draft on the same topic

Completion of the Error Checklists

Each time students received their first or second draft, they had to complete the error checklist of that particular draft (see Appendices E and F for the Error Checklist and the example of how to fill out the list). To fill out the checklist, students had to first count the errors they made in each category and then fill the number in each cell. The total number of errors in each major category could be computed by adding all the numbers of its subcategories together. The total number of errors made in each draft could also be calculated by adding the number of each cell in the same column. By dividing the total number of errors by the total number of words written, an error rate of each draft was obtained. The administration of the checklist was essential because it allowed students to see if they improved their writing accuracy. Also, it allowed them to diagnose their own weakness in grammar.

Students of the two treatment groups however did not weigh the task of filling out their error checklist equally. Completion of the checklist was obviously an easier job for students in the code correction group than their counterparts in the direct correction group. They had only to categorize the errors they made according to the codes used by their teacher to address these errors. To complete the checklist, they merely counted the errors of each category and filled out the form

The job of completing the checklist demanded greater cognitive efforts from students in the direct group because they had to identify the types of errors. Lee (1997) attributed students' failure in error correction to their inability of detecting their errors. Even for students in the direct group who received correction answers from their teacher, the job of categorizing the errors could still be difficult for students of lower language proficiency due to their deficit in grammar knowledge.

Implementation of the Attitude Questionnaire

At the end of the experiment when students received their second draft on the fourth topic, the attitude questionnaire was administered to investigate their attitudes and perceptions of the treatment they received (see Appendix C and D for the attitude questionnaire).

Instruments

The instruments implemented in the present study included, first, JCEE rating scale, an attitude questionnaire, a taxonomy of error codes and an error checklist,

Rating Scale

The grading of all the nine compositions in the present study was based on a holistic rating scale used in Joint College Entrance Examination (JCEE) (see Appendix H). The scale includes five components: content (5 points), organization (5 points), grammar (4 points), vocabulary (4 points), and mechanics (2 points). According to this scoring guide, the full score is 20, and the compositions graded are approximately divided into five levels of proficiency: Excellent to Very Good (19-20 points), Good to Average (15-18 points), Average (10-14 points), Mediocre to Poor (5-9 points), Very Poor (0-4 points).

This scoring guide is chosen mainly due to its popularity among Taiwanese high school teachers in the local context. Besides, its provision of clear guidelines in scoring compositions helps ensure a higher inter-rater reliability.

Students' Attitude Questionnaire

The questionnaire administered in this experiment made reference to the questionnaires used in Ferris' (1995) (see Appendix C and D for the attitude questionnaire). The devise of the questionnaire was to explore students' attitudes toward error correction. Item 1 to Item 3 were Liker-scale items which investigated students' self perceptions of their improvement in writing ability, self editing ability

and writing autonomy respectively. Item 4 and 5 were multiple choice questions. Item 4 surveyed students' perceptions of correction responsibility while Item 5 explored the students' preference for error correction. The last item of the questionnaire was an open-ended question designed to elicit further information concerning their perceptions, and attitudes toward error correction.

Table of Error Codes

The design of the Table of Error Correction used for decoding error symbols in code correction was mainly adapted from Ferris et al.'s (2001) study. I also made reference to the categorization by other researchers (Chandler, 2003; Huang, 1988; Lalande, 1982; Lee, 2004). This table appeared in two forms (see Appendix B for both forms of the table). The first one displayed an original version with clear examples offered while the second one was a simplified version with error symbols arranged alphabetically for students' quick reference.

In this table, errors were classified into five major categories— each of which was further divided into 2 to 6 subcategories. The main categories did not appear on students' draft. The grouping of the subcategories into the main categories was only for students' convenience to detect their weakness in grammar. What was utilized in the code error correction was the subcategories, each of which stood for an error symbol. For students' quick decoding of the error symbols they received in their feedback, these 19 symbols of subcategories were rearranged alphabetically to form a simplified version of the Table of Error Codes.

The five major categories of this table included “verb errors”, “mechanics errors”, “wrong form”, “wrong usage”, and “sentence structure errors.” “Verb errors” included all errors in verb tense or form, and the subject-verb agreement errors were also included. The categories of “wrong form” and “wrong usage” both addressed errors occurring within phrase or word boundaries. The former type focused more on the

errors of inflectional endings including incorrect plural or possessive ending; whereas the latter one emphasized the specific lexical errors in word choice such as preposition, noun and article errors, or the collocation problems. The spelling errors were included in this category only when the misspelling did not result from the wrong inflection of word endings. As to the category of “sentence structure errors”, it contained errors in sentence or clause boundaries such as the occurrence of incomplete sentences or fragments, wrong word order, and the improper application of conjunctions or sentence connectors such as “however” and “moreover.

The four major categories introduced previously were derived from Ferris’ classification of errors. However, the remaining type of errors was different between the researcher’s design and Ferris’. Ferris grouped the article and determiner errors into a major category. The researcher however included these errors in the major category of “wrong usage.” Instead, she listed another major category of “mechanics errors” which were also among common errors made by her own students.

Error Checklist

The error checklist was designed according to the Table of Error Codes (see Appendix E and F for the checklist). Each empty cell in the checklist represented the number of that particular type of errors made in a draft. By counting the number of errors to complete their checklists, students in both groups were able to diagnose their weakness in grammar and have their grammar awareness raised. They could also observe their improvement or deterioration in their written accuracy from the change in the error rate of each draft.

Data Sources and Analyses

The data collected for this study came from the following sources: (1) the overall writing score of the writing pretest (the first draft on the first topic) and the writing posttest (the first draft on the fourth topic), (2) the error rate of the pretest writing and

the posttest writing and (3) the responses to the attitude questionnaire. For all the quantitative data collected from the first three parts including the writing scores, error rates as well as the responses to the first five questions in the questionnaire, a statistical program SPSS 12.0 for Windows is applied for analysis. The data collected from the last question in the questionnaire is analyzed qualitatively for comparison and discussion.

Pretest and Posttest Writing Scores

The writing pretest and posttest were the first drafts written on the first and fourth topic. They were graded by two raters using the Joint College Entrance Examination (JCEE) scoring criterion (see Appendix G for the rating scale). The two raters were the researcher herself and a second rater, a teacher with five-year teaching experience in teaching English writing to high school students. Before the experiment began, the two raters participated in a training session in which they marked several samples of students' essays. After both raters graded all the pretest and posttest compositions which were mixed and numbered to disguise the writer's identity, the Pearson product-moment correlation coefficients were employed to calculate the inter-rater reliability between the two raters. The inter-rater reliability was .86 for the pretest writing and .89 for the posttest writing at a .05 significant level. The average of the two individual scores given by the two raters respectively was used as the final overall composition score which would be further subjected to SPSS for statistical analyses. The statistical procedures used to analyze the collected quantitative data of the overall writing scores included descriptives for the means and the standard deviations, MANOVAs which assessed differences between two different groups at two levels of proficiency, and two-way ANOVAs which examined the interaction between treatment types and writing proficiency levels.

Pretest and Posttest Writing Error Rates

To measure students' improvement or regression in their writing accuracy, the "error rate", or the relative frequency of errors, was used for data analysis. To calculate the error rate in each draft, the total number of errors was divided by the total number of words in each written text. The formula was as follows: Error Rate = Number of total occurrences of errors / Number of total words written. The number of errors was counted by both the researcher and the student writers themselves. All the participants counted the errors they made when they were required to fill out their error checklists and the results were double-checked by the researcher.

To examine the improvement in student writing accuracy, error rates of the pretest and posttest writing were subjected to SPSS for analysis. The statistical procedures used for analyses included descriptives for the means and the standard deviations, MANOVAs and two- way ANOVAs. These analyses tested the within-group and across-group differences to reveal the impact of different writing proficiency on the effects of the treatments.

Responses to Attitude Questionnaire

To analyze the quantitative data collected from the Liker-scale and multiple response items, statistical procedures including descriptives (frequency and percentage) and a chi-square test (to examine the percentage differences across different groups) were applied. As to the data collected from the last open-ended question, it was analyzed qualitatively by selecting some excerpts from students' answers for comparison and discussion, so that students' opinions about the two treatment types were revealed.