ผลของวิธีสอนแบบวิเคราะห์รูปแบบงานเขียนเชิงอภิปรายโวหาร
ต่องานเขียนของผู้เรียนภาษาอังกฤษเป็นภาษาต่างประเทศ

นายวิวรรธน์ อุดมยมกกุล

A GENRE-BASED APPROACH TO TEACHING ARGUMENTATIVE WRITING : EFFECTS ON EFL STUDENTS' WRITING PERFORMANCE

Mr. Wiwat Udomyamokkul

A GENRE-BASED APPROACH TO TEACHING

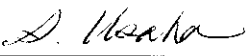# ARGUMENTATIVE WRITING : EFFECTS ON EFL STUDENTS'

WRITING PERFORMANCE

Suranaree University of Technology has approved this thesis submitted in partial fulfillment of the requirements for a Master's degree.

Thesis Examining Committee

_____

(Dr.Sarit Srikhao)
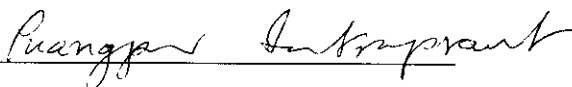
Chairperson

_____

(Asst. Prof. Dr.Siriluck Usaha)

Member (Thesis Advisor)

_____

(Assoc. Prof. Dr.Puangpen Intaraprawat)
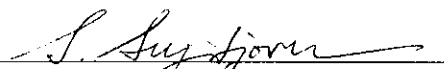
Member

_____

(Assoc. Prof. Dr.Thai Tipsuwannakul)
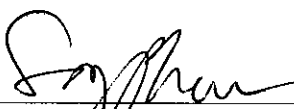
Member

_____

(Dr.Jiraporn Sangarun)

Member

_____
(Assoc. Prof. Dr.Sarawut Sujitjorn)
Vice Rector for Academic Affairs

_____
(Assoc. Prof. Songporn Tajaroensuk)
Dean of Institute of Social Technology

วิวรรธน์ อุดมยมกกุล : ผลของวิธีสอนแบบวิเคราะห์รูปแบบงานเขียนเชิงอภิปรายโวหารต่อ
งานเขียนของผู้เรียนภาษาอังกฤษเป็นภาษาต่างประเทศ (A GENRE-BASED
APPROACH TO TEACHING ARGUMENTATIVE WRITING : EFFECTS
ON EFL STUDENTS' WRITING PERFORMANCE) อาจารย์ที่ปรึกษา : ผู้ช่วย
ศาสตราจารย์ ดร. ศิริลักษณ์ อุสาหะ, 160 หน้า. ISBN 974-533-372-7

งานวิจัยกึ่งทดลองเรื่องผลของวิธีสอนแบบวิเคราะห์รูปแบบต่องานเขียนเชิงอภิปรายโวหาร
ของผู้เรียนภาษาอังกฤษเป็นภาษาต่างประเทศมีวัตถุประสงค์เพื่อศึกษาและวิเคราะห์ประสิทธิภาพ
ของวิธีสอนเขียนแบบวิเคราะห์รูปแบบต่อผลสัมฤทธิ์ทางการเรียนการเขียนเชิงอภิปรายโวหารของ
นักศึกษากลุ่มทดลอง โดยเปรียบเทียบกับกลุ่มควบคุมซึ่งเน้นวิธีการเรียนการสอนแบบเน้นกระบวน
การเขียนเพียงอย่างเดียว กลุ่มตัวอย่างที่ใช้ในการวิจัยคือ นักศึกษามหาวิทยาลัยเทคโนโลยีสุรนารีที่
ลงทะเบียนเรียนวิชาการเขียนภาษาอังกฤษ 5 ในภาคการศึกษาที่ 2 ของปีการศึกษา 2546 จำนวน 55
คน โดยแบ่งเป็นกลุ่มทดลอง 28 คนและกลุ่มควบคุม 27 คน ข้อมูลที่ใช้ในการวิจัยคือ เรียงความ
ฉบับร่างแรกและร่างสุดท้ายของนักศึกษาทั้งสองกลุ่ม

ผลการวิจัยที่สำคัญพบว่า การสอนแบบวิเคราะห์รูปแบบสามารถช่วยให้ผู้เรียนเข้าใจได้เร็ว
และเขียนได้อย่างมีประสิทธิภาพถูกต้องและเหมาะสมกับคุณลักษณะของงานเขียนเชิงอภิปราย
โวหาร โดยที่คะแนนงานเขียนเรียงความเชิงอภิปรายโวหารฉบับร่างแรกของกลุ่มทดลองที่ได้จากผู้
ประเมินแตกต่างและมากกว่ากลุ่มควบคุมอย่างมีนัยสำคัญทางสถิติที่ระดับ .05 อย่างไรก็ดี คะแนน
งานเขียนฉบับร่างสุดท้ายที่ได้จากผู้ประเมิน พบว่าทั้งสองกลุ่มไม่แตกต่างอย่างมีนัยสำคัญที่ระดับ
.05 ซึ่งชี้ให้เห็นถึงผลสัมฤทธิ์ทางการเรียนในระยะยาวของวิธีการสอนแบบเน้นกระบวนการเขียน
นอกจากนี้ ผลสำรวจความคิดเห็นของนักศึกษากลุ่มทดลองที่ได้จากแบบสอบถามชี้ให้เห็นถึง
ประโยชน์ของการสอนแบบจำแนกรูปแบบที่มีส่วนช่วยในการนำเสนอและเรียบเรียงความคิดและ
เนื้อหาในการเขียนเรียงความเชิงอภิปรายโวหารของผู้เรียนกลุ่มทดลอง

สาขาวิชาภาษาอังกฤษ   ลายมือชื่อนักศึกษา_____

ปีการศึกษา 2547   ลายมือชื่ออาจารย์ที่ปรึกษา_____

                                                              ลายมือชื่ออาจารย์ที่ปรึกษาร่วม_____

                                                              ลายมือชื่ออาจารย์ที่ปรึกษาร่วม_____

WIWAT UDOMYAMOKKUL : A GENRE-BASED APPROACH TO

TEACHING ARGUMENTATIVE WRITING : EFFECTS ON EFL

STUDENTS' WRITING PERFORMANCE. THESIS ADVISOR : ASST.

PROF. SIRILUCK USAHA, Ph.D. 160 PP. ISBN 974-533-372-7

This posttest-only quasi experiment using control group investigated whether the use of genre-based approach including explicit instruction of rhetorical patterns of English argumentative discourse was effective in helping students gain control of academic argumentative genre in comparison with the effects of control treatment which focused relatively more on the teaching of writing process.

Fifty-five non-randomized Thai undergraduates of Suranaree University of Technology participated in the study; 28 of whom as the experimental group were exposed to the genre-based instruction and the other 27 as the control group were exposed to the instruction on the process of writing alone. Data used for the analysis were the first and final drafts of 55 argumentative essays. The scripts were read and scored by raters for the overall impression of the writing (holistic scoring) and for the three separate Toulmin-based persuasiveness qualities including claims, reasons, and rebuttals to counter-arguments (analytic scoring).

The resulting outcomes of the study were twofold. First, the experimental group outperformed the control group as indicated by the significantly higher gain scores awarded on the first drafts' development and organization of claims and rebuttals to counter-arguments. Second, given the opportunities for multiple drafting, revising, and editing, both groups were equivalent in their final-draft performance as indicated by the holistic gain scores. Responses to the questionnaire items suggested the usefulness of the genre-based

instruction experienced by the experimental students in helping them effectively develop and organize their ideas in their essays.

School of English   Student's Signature_____

Academic Year 2004   Advisor's Signature_____

                                                      Co-advisor's Signature_____

                                                      Co-advisor's Signature_____

ACKNOWLEDGMENTS

TABLE OF CONTENTS

Page

CHAPTER

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Background and Rationale

In recent years, composition practitioners may have seen increased emphasis being placed on the notion of genre and genre-based pedagogy in English language teaching particularly in the teaching of EAP/ESP writing. Such a movement is evident in a recent proliferation of genre conferences and applications in North America and Australia where genre theory and its applications have been better developed than elsewhere. In Australia, for example, genre-based approaches have been applied in academic writing, English in the workplace, adult second language literacy development, and language development in schools (Paltridge, 1996).

The concepts and implications of genre, however, are not new. From the late 1980s, genre has been a popular framework for analyzing the form and function of nonliterary written discourse or "an examination of the constituent parts or move structures of text which represent the writer's communicative purpose" (Flowerdew, 2000, p. 369). According to Hasan (1989), genre can be defined as:

> a text, either spoken or written, that serves a particular purpose in a society and is composed of a series of segments, called moves. Some of the moves in a genre are obligatory in that they are necessary to achieve the communicative purpose of the genre, whereas others are optional or those which speakers or writers may choose to employ if they decide those moves add to the effectiveness of the communication but do not alter the purpose of the text. (quoted in Henry and Roseberry, 1998, p. 147)

Thus, the work on genre analysis is to identify how these moves are organized in a given genre or text type (schematic/rhetorical structures of texts) and its linguistic features used to realize the communicative purpose of that genre. The outcome of genre studies has resulted in many teaching materials for professional and academic purposes. John Swales and Christine Feak's *Academic Writing for Graduate Students: Essential Tasks and Skills* (1994), Robert Weissberg and Suzanne Buker's *Writing Up Research: Experimental Report Writing for Students of English* (1990), and Tony Dudley-Evans's *Writing Laboratory Reports* (1985) are, among others, some examples of the applications of the results of genre analysis in the teaching of written genres in the language learning classroom.

With respect to the teaching of genres in writing, the debate has developed over the efficacy of such instruction. Two opposite views can be contrasted where the proponents of genre-based approach see the necessity of explicit instruction of genres as writing, like other skills, can be learned at least in part through direct guidance (Williams and Colomb, 1993), whereas some composition researchers and scholars cast doubt against such teaching, viewing that explicit genre instruction has only restricted value in improving students' writing and is generally unnecessary because students acquire genre knowledge tacitly (e.g., Freedman and Medway, 1994; and Sawyer and Watson, 1987). To negotiate conflicting views, Badger and White (2000) proposes the efficacy of what they called a process-genre approach to teaching writing informed by product, process, and genre views of writing and writing development:

> The writing class recognizes that writing involves knowledge about language (as in product and genre approaches), knowledge of the context in which writing happens and especially the purpose for the writing (as in genre approaches), and skills in using language (as in process approaches) …writing development happens by drawing out the learners' potential (as in process

approaches) and providing input to which the learners respond (as in product and genre approaches). (pp. 157-158)

In higher educational settings, the role of English language is important as it has acquired the status of an international lingua franca in almost any field of study. "The ability to write a good essay as a major vehicle of individual expression often exclusively determines a student's success in his or her area of study" (Mitchell, 1994, p. 82). Amidst the enthusiasm over genre as a teaching tool, few studies have evaluated the effects of genre-based pedagogy on students' language and literacy development. O'Brien (1995) has observed that while educators believe that student essays constitute an important learning tool in the development of conceptual understanding and linguistic expression, the essays, however, are "rarely studied in any details by text linguists" (p. 442). As noted by Henry and Roseberry (1998), "the arguments for and against the genre approach in ESP/EAP have been limited to the theoretical, and few, if any, attempts have been made to evaluate the approach empirically in an ESP/EAP context" (p. 148). This is affirmed by Dudley-Evans (2001) that research into the discourse structure of assignments or essays is limited. This lack of assessment of effects of genre-based pedagogy on students' acquisition of language skills particularly reading and writing has perhaps been one reason why genre-based teaching is sometimes controversial (Freedman, 1994).

Although there have been few studies on genre and its applications so far and their resulting generalizations may be viewed as anecdotal, empirical studies consistently reported positive effects of genre-based teaching for non-native English speakers though some limitations were suggested as well (Hyon, 2001; Henry and Roseberry, 1998; Johns, 1999; Mustafa, 1995). In an attempt to investigate language

teachers' perspectives toward genre in writing curricula, Kay and Dudley-Evans (1998) have observed that of 48 teachers participated in the study, almost all of them hold mixed views on genre in teaching writing, and some believe a genre-based approach helps students "discover how writers organize texts" (p. 310). However, some instructors raise concern over the prescriptiveness and demotivating nature of strictly genre-based teaching. Nevertheless, the similar positive gains suggested by the studies cited above might lead to the generalizations that a teaching approach focusing on rhetorical organization (explicit genre instruction) can be successful in an EAP/ESP teaching situation with reasonably advanced learners (Henry and Roseberry, 1998), in raising students' awareness of genre (term paper) conventions more often than those who had not received the genre instruction (Mustafa, 1995), and in exploiting the benefits of using genre as a socioliterate approach to teaching reading and writing in developmental university composition for language minority students (Johns, 1999).

Based on the data retrieved from the Thai Thesis database, there have not been many research studies on the practice of genre-based approach to teaching literacy to Thai EFL learners especially to adult learners. A number of studies carried out in the Thai context seems to provide results similar to those cited previously as shown by positive learning outcomes of students receiving genre-based instruction. However, most of the instructional studies reviewed by the researcher took place in secondary school contexts mostly covering the skill areas of reading, listening, and speaking with less undertaken on the writing skill (Thai Thesis CD-ROM, 2003). For instance, Likhasith's (1995) investigation of the effects of using a genre-based approach in teaching reading on students' development in story reading and writing skills shows

that the students receiving genre-based instruction gain significantly higher scores in story writing tests than those received by the control group. Similarly, Cheedlamduan's (1995) experimental study reveals favourable results of the use of genre-based approach in teaching reading and writing to Thai secondary (Mathayom Suksa II) students as reflected by a significant difference in scores on the English Reading and Writing Ability Test between the experimental and control groups. To the best of the author's knowledge, research into the use of process genre-based approach and its effectiveness in the teaching of EFL writing in post-secondary educational contexts rarely exists.

Amidst continued development in genre-based teaching applications and suspicions over its efficacy, composition scholars and researchers call for the need to examine the effects of genre instruction on students' reading and writing development. While questioning the merits of the explicit nature of genre pedagogy, Freedman (1993) has adopted the fact that the research evidence concerning genre acquisition is limited and called for more focused research and theoretic consideration of the issue. Similarly, Hyon (2001) suggests that as approaches to genre-based pedagogy vary significantly, investigations need to be conducted of different approaches to such teaching in a variety of classroom contexts. In addition, in stating the limitations of their study, Henry and Roseberry (1998) raise a concern on another research direction on learning and transfer for determining the extent to which students can transfer the skills acquired during genre-based instruction to different activity types.

In an attempt to address the research need and ascertain the claims forwarded by genre theory and its pedagogical applications, it is interesting to investigate how effective a genre-based approach is in the teaching of EAP writing to a group of Thai

undergraduates, expecting that the approach will be facilitative in helping EFL learners to gain control of the argumentative academic genre in the topics relevant to their areas of study. Hopefully, the learners' writing performance as a result of the genre-based instruction will represent a positive backwash encouraging composition teachers to adopt practices that are in line with the current best thinking in the field with respect to pedagogy.

1.2 Statement of the Problem

In terms of language skills, writing in English as a foreign language is perceived to be a Herculean task by many if not most EFL students. This comes as no surprise that apart of a lack of sufficient language knowledge, the difficulty faced by many of these students is largely attributed to the lack of composing skills and strategies necessary to accomplish the specifically assigned writing tasks.

For learners in advanced academic settings like in the university, the use of writing extends beyond the basic goals for personal expression or presenting school-based information. Rather, the university student writers are often expected to analyze and interpret information critically, synthesize disparate sets of information, argue alternative perspectives, and create and present information through effective writing. The ability to think and write critically is often linked with argumentation and the writing of argumentative essays. This critical thinking is often assessed by asking students to identify an issue, consider different views, take and support a stand, and reflect and respond to competing viewpoints.

Given the presumed importance of argument and that competence in persuasive discourse is needed by university students (Mitchell, 1994; Grabe and

Kaplan, 1996), several sources indicate that students still have a great deal of difficulty writing argumentative discourse, and so display so little skill in argumentation. Research into students' argumentation skills and writing reveals highly unsatisfactory student writing performance that falls short of expected valued-way of academic writing. A large-scale longitudinal L1 writing survey called the National Assessment of Educational Progress (NAEP) reveals a dismal writing performance of American students, suggesting that no American eighth-graders were able to write an elaborate persuasive essay that took a stand and provided a cohesive, detailed discussion of the issue and alternatives. The survey also indicates that despite four additional years of schooling, less than two percent of twelfth-graders were able to write elaborate essays (Applebee, Langer, Mullis, Latham, and Gentile, 1994).

Some years ago, a body of research into persuasive student texts has identified some common weaknesses in native and non-native speaker student writing performance. One weakness is inadequate content including insufficient information to support their viewpoints, ineffective formulation of claim or thesis, lack of awareness of audience, and lack of clarity and clear focus (e.g., Pringle and Freedman, 1985; Applebee, Langer, and Mullis, 1986; Crowhurst, 1987, 1991). A second weakness is poor organization, which frequently results in students' production of incohesive and incoherent texts (e.g., Freedman and Pringle, 1984; White, 1989; Hyland, 1990). In addition, students have been found to write narration when asked to persuade (Wilkinson, Barnsley, Hanna, and Swan, 1980; Crowhurst, 1983a; White, 1989; Reppen, 1995).

Empirical studies suggest several causes accounting for such students' poor performance in writing arguments. One important explanation is that students are

more acquainted with narration and are likely to produce a writer-based prose that merely reflects and reproduces the contents of subject matters they know or remember. Persuasive writing as a reader-based prose (as opposed to a writer-based prose) takes the reader into account. As Flower notes, composing a writer-based prose is a highly cognitive demanding task as it entails "selecting a focus of mutual interest, moving from facts and details to concepts, and changing narrative or episodic style into expository style organized around the writer's purpose" (1979, p. 37). Bereiter and Scadamalia (1982) argue that students operate from oral discourse schemata when faced with a writing task. They believe that young students have not developed appropriate written discourse schemata to enable them to write persuasive discourse successfully. Stein and Glenn (1979) have shown that learners do learn to use narrative forms successfully and fairly early, explaining that learners learn to use narrative in the course of daily communication. They also note that the features of formal argument might not be learned as readily from daily oral interchanges.

Another source of explanatory evidence from the work in the area of contrastive rhetoric reveals students' L1 rhetorical predisposition in approaching the writing task which is often associated with the interference of differing L1 rhetorical patterns when writing in English (Kaplan, 1966; Connor and Kaplan, 1987; Hinds, 1990; Connor, 1996). An analysis of expository texts written in Japanese, Korean, Chinese, and Thai, a group of east Asian countries where English is used only in business and higher education, reveals individual countries' preferred rhetorical modes of written expressions that differs in some ways from English conventions (Ostler, 1987; Hinds, 1987, 1990). ESL composition researchers and teachers like Barbara Kroll, Dana Ferris, and Tony Silva have come to similar conclusions that ESL student

writers frequently bring to the academic writing task both linguistic and rhetorical deficits (Kroll, 1990; Silva, 1993; Ferris, 1994).

Some researchers have argued that the reason why students seem to display so little skill in persuasive writing is just that they simply have not had enough exposure to or instruction in strategies for persuasive writing. The fact that formal instruction in argumentation and the writing of argumentative discourse which is typically delayed until higher secondary grades (eleventh grade as in L1 writing situations, as Yeh [1998] argues) probably represents a cause of lack in instructional studies in the domain of argumentation. This might be a reason why most research of the late 1980s and early 1990s have been largely focusing more on the analysis of student persuasive texts than on the extent to which instruction can affect the quality of their written arguments. According to McCann (1989), instruction in writing of any kind in secondary schools is slight, observing that even college composition courses reserve argument for the end of the course. Students write infrequently in secondary schools and that instruction in writing is minimal (Applebee et al., 1994).

While it can be seen that student argumentative writing has been widely studied by rhetoricians and scholars, the earlier work has thus far been restricted to students who were writing in their first language or to international students of advanced second language proficiency who were studying in English speaking countries for many years particularly in North America. In contrast, little research has been conducted in evaluating the effects of different approaches for teaching argumentative writing at different grade levels as well as in post-secondary levels (Fulkerson, 1996). Michael Long and Jack Richards have suggested that while much is known about how second language writers write and learn to write, much remains to

be discovered about the kinds of writing instruction student writers are most likely to benefit from (cited in Kroll, 1990, p. viii). As noted in the previous section, based on a review of database on relevant studies using EFL Thai students as subjects, research into student argumentative writing is rare and has been largely confined to the linguistic analysis of argumentative texts written by English major undergraduates with few of these studies attempting to measure and compare the effects of using different approaches in teaching students argumentative writing (TIAC, 2003).

At Suranaree University of Technology (SUT), the undergraduate students are required to enroll in an English course on argumentative writing known as English 5, which is the last compulsory English module offered. Based on the questionnaire responses, the student subjects of the study declare that they have never been taught to write argumentative prose at all before entering the university, though they might have had engaged in writing activities somewhat similar to persuasion during high schools but not aware of it. In addition, they admit that they also have problems within sentences in terms of vocabulary and grammar. However, it is above the sentence level that poses the challenge for them particularly when they are required to create whole texts such as argumentative essays. Generally, among other skills of English language taught in Thailand in the formal schooling system, writing is viewed to be the most disregarded and perhaps the least emphasized (Pimsarn, 1987; Nipitkul, 1995; Chinnawong, 2002). More specifically, a recent study on the writing performance of Thai science undergraduates reveals that grammar and vocabulary is not only the two major problematic areas that must be attended to but also discourse organization and idea development (Chinnawong, 2002). In addition, most Thai

students nearly always find it difficult to write in English because they are anxious about what to write and how to start writing (Pimsarn, 1987).

While native-speaker student writers still have problems producing argumentative written discourse because of several causes mentioned above, the difficulty experienced by students from non-English speaking backgrounds is highly likely to be more greater. This is mainly because of minimal (or lack of) formal instruction in necessary writing skills in general or in argumentation in particular and of limited exposure to standard written persuasion in English, let alone a deficient language knowledge factor. As L1 educators have claimed, formal instruction in argumentation and the writing of arguments is usually delayed until higher secondary grades or college levels in the L1 situations. In contrast, it is obvious that writing instruction in such area is not available at all in certain L2 settings at least as pointed out by the student subjects above. Hence, EFL student writers like those participated in the study are traditionally under-prepared for performing academic writing tasks effectively.

In the context of this study and in the composition classroom elsewhere, the use of process approach is undoubtedly useful in helping students overcome their emotional as well as linguistic barriers to writing. By directing students to follow the recursive writing process including prewriting, drafting, revising, and editing, teachers hopefully expect students to become engaged in strategies and activities that skilled writers usually do in producing an effective prose. In the teaching of argumentative writing, an emphasis on purpose and audience, the two aspects of social context in which students learn to write, is useful in making students aware of the social and communicative purpose of writing. An awareness of these two aspects of writing

context enters the writing process when student writers generate their ideas during prewriting, decide on how those ideas are communicated during drafting, and ensure that the information is clear during revising and editing (Raphael, Englert, and Kirschner, 1989). However, an observation of student writing (drafts) points out that many students still have trouble developing and organizing their ideas when preparing early drafts and many of them have to go through several drafts before they can create an acceptable one.

Recalling that L2 student writers, unlike L1 students, usually lack implicit knowledge of rhetorical pattern and organization of genre, an instructional approach focusing on contexts of writing then may not be sufficient for these students. Although students are encouraged to write and revise several times, attempts to work from student texts toward the genre may not be an economic and ideal way in enabling them to write. Learning to write in this way becomes a matter of several corrections and rewriting according to teacher feedback with students relying less on their own knowledge and performance as a result of instruction. Taking this into account, students' lack of genre knowledge with respect to its structural pattern and organization calls for a purposeful intervention by teacher, and this could be done as early as in prewriting, even before they start to write.

Based on research on genre and its applications as discussed in the previous section, the so-called genre-based instruction focusing on explicit teaching of genre knowledge with respect to its forms and functions at a level of whole text then becomes relevant in the situation of this study. Because of lack of exposure of formal written persuasion in English, students need to be sensitized to genre form by reading, discussing, and analyzing examples of it. Explicit knowledge of genre form including

its rhetorical pattern and organization is expected to work in tandem with students' knowledge of contexts of their writing and to help them fulfil the informational as well as organizational demands of the assigned genre. This kind of knowledge works in such as way that while an awareness of purpose and audience has a major impact on the ideas generated and how those ideas are communicated, explicit knowledge of rhetorical structure and organization affects students in terms of how texts should be structured and organized in certain ways to fulfil its communicative purposes. The expected usefulness of the process-genre approach utilized in this study therefore is that while students are aware of the information they need to express with reader in mind, they also become capable of knowing how to present their ideas in writing in a sufficiently clear and organized way.

The discussion so far does not mean to question the pedagogic potential of process approach but intends to draw attention to how the approach should be adapted to create an optimal learning condition for EFL developmental writers. For students with sophisticated language proficiency, an understanding of argumentation principles and composing process is likely to be adequate in facilitating them to produce argumentative prose as they may have some ideas of what a piece of written persuasion is from their educational background and knowledge of language. For learners with limited language proficiency especially in writing skills, they would be disadvantaged because without being made clear to them the characteristic textual features of the writing task, they cannot be expected to know what it means to write clear, focused, well-organized and coherent prose texts. The position of this instructional study then is that students will profit from both types of instruction. In other words, the use of integrated process-genre approaches to argumentative writing

instruction is expected to help students gain control of the assignment genre more quickly and also benefit those who are likely to fail.

As a quasi-experiment, the intent of this study is to examine hypothesized effectiveness of utilizing a process-genre approach in SUT undergraduates' argumentative writing instruction using the ad hoc criteria in assessing the quality of student argumentative writing. Consistent with the process view of writing and the development of writing ability and the pedagogical claims forwarded by genre pedagogy, prewriting instruction incorporating explicit instruction in the domain of argumentation is expected to help students approximate control of the argumentative genre more quickly and effectively. Students with a relatively well-developed knowledge of composing process and of textual features of argumentative written discourse should be able to perform better on the writing tasks related to argumentation.

1.3 Purposes of the Study

The major aim of the study is to investigate the effectiveness of genre-based instruction in accommodating the learners to produce an effective argumentative essay. In addition, the other objective is to see whether the students exposed to the genre-based instruction find that their developed abilities to write an argumentative essay are actually facilitated by the instruction.

1.4 Research Questions and Hypotheses

The questions to be investigated in the study are:

1. Will genre-based instruction facilitate the student writers to write an effective argumentative essay?

2. Will the students find the instruction useful in helping them to produce an effective written argumentation?

Two relevant hypotheses to be tested are:

1. Genre-based instruction will enhance the experimental subjects' ability to produce an argumentative text that incorporates salient features of argumentation including an effective development and organization of claims, reasons, and rebuttals to counter-arguments in their scripts in comparison to the control subjects' writing performance.

2. Students exposed to the genre-based instruction will provide a positive feedback toward the instruction.

1.5 Significance of the Study

The findings of the present study could have a beneficial backwash as follows:

1. It may provide useful pedagogical implications by presenting the advantages that can be realizable from using a genre-based instruction in the teaching of writing particularly in the prewriting stage or in helping the students write up the essay.

2. The results of the study will portray what the students perceived to be the advantages of the instruction and the difficulties they encountered when writing up an argumentative essay, be it rhetorical or language or both.

3. The performance assessment outcome of this research will allow the researcher and the reader alike to acknowledge a range of students' writing performance and the characteristics of argumentative writing produced by students as classified by the rating scales, as a result of the two different types of argumentative writing instructions used in this study.

1.6 Assumptions

At the outset of this research, the followings are presumed:

1. The learning outcomes as shown by the final-essay scores gained by both groups of students participated in the study as a result of the teaching of process writing utilized by both classes' teachers are hypothesized to be not different in a meaningful way, and, if there exists any difference, such difference is insignificant.

2. The impact of the hypothetical instruction is hypothesized to be large in the first stage of the writing process (prewriting/drafting), compared with that of the comparison group.

3. The effect of the genre-based instruction on the learning outcome may not necessarily extend to other areas such as language use other than the acquisition of rhetorical competence, though in some cases (e.g. for some relatively advanced students compared with peers), it could possibly do.

1.7 Scope and Limitations of the Study

1. The study takes place within the tertiary educational settings of Suranaree University of Technology (SUT), Nakorn-Radchasima, Thailand. The subjects participated in the study are 55 Thai undergraduates from varied fields of science and engineering studies enrolled in English 5, an argumentative writing course, at SUT in Trimester 2 of the academic year 2003.

2. Genre-based instructional lessons and materials are devised by the teacher who practices a genre-based approach and for the genre class students only.

3. The research intent is to determine the impact of hypothetical instruction by evaluating the students' argumentative writing performance in three areas of rhetorical qualities of effective argumentation, that is, the development and organization of content including (1) claims, (2) reasons, and (3) rebuttals to counter-arguments.

4. Three out of the original six elements of Stephen Toulmin's (1958) framework for argument analysis are simplified and adapted to construct the rating scales for analyzing the students' argumentative essays.

1.8 Expected Outcomes

Corresponding to the purposes of the study, the following results are expected:

1. Genre-based instruction will have a positive effect in facilitating the learners to acquire argumentative writing skill as shown by the scale scores awarded by raters on the essays using the specialized constructed rating scales.

2. The rating scales (holistic and analytic) constructed for assessing the students' argumentative essays will prove to be valid and reliable research instruments

as shown by a set of computed values of test statistics used for measuring scorer reliability.

3. The study will provide insights into the practice of genre-based pedagogy in terms of its effectiveness in facilitating the subjects to gain control of the argumentative genre.

4. The students will find the genre-based instruction useful in helping them to learn to write an effective argumentative essay as expected by the course.

1.9 Definitions of Terms

Unless otherwise stated, the following terms used in the study assume specific meanings as explained below:

1. "Genre-based approach/instruction/treatment" refers to an approach to teaching argumentative writing that emphasizes the use of direct instruction of rhetorical elements and organization of argumentative genre by having the students analyze the form and function of text.

2. "Non-genre approach" means an approach to teaching argumentative writing that relies relatively more heavily on the teaching of abstractions of argumentation through the writing process encouraging multiple drafting, revisions, with feedback from the teacher during each stage of the writing process.

3. "Rating scale(s)" means the rubric constructed as a research tool to be used by raters to scores the students' essays.

4. "Argumentative essays" are sometimes referred to as "persuasive essays." Both terms are similar in meanings and often used interchangeably in the

writing literature. An argumentative or persuasive essay means an essay that argues or persuades by taking a position of a controversial issue, giving reasons to support the position, and addressing and refuting possible counter-arguments.

5. "EFL" stands for English as a Foreign Language and means the teaching and learning of English in communities where it is not widely used for communication. "EFL students" refer to students who learn English in Thailand or countries where English is not used widely for communication.

CHAPTER 2

LITERATURE REVIEW

This chapter consists of three sections. The first section describes argument with respect to its definition, structure, and organization. The second section provides a review of literature on the approaches to teaching writing and how each of them is complementary to one another. The third and final section focuses on the related issues on assessing L2 writing with an emphasis on the rating of argumentative writing.

2.1 Argument: Definition, Structure, and Organization

2.1.1 Defining Argumentation-Persuasion

Both these two terms *argumentation* and *persuasion* can be found to be used interchangeably to carry quite the same meanings in the writing literature. One instance of this was claimed to be the replacement of 'persuasion' with 'argument' in the writing curriculum in England and the United States in the eighteenth century. Prior to the change, the curriculum was dominated by a four-component model of discourse: description, narration, exposition, and persuasion (Connor, 1996). Another instance was that, as McCann has noticed, a ten-year study of student achievement in writing sponsored by the National Assessment of Educational Progress (NAEP) asked students to write arguments in response to persuasive topics (1989). Nevertheless, it was not until Kinneavy's classification of discourse in 1971 that persuasion was again

included as one of the four major aims of communication: referential, persuasive, literary, and expressive (Connor & Lauer, 1988).

As the field of modern communication as well as academic argumentation has been much influenced by Aristotelian rhetoric as a preferred mode of inquiry and persuasion, argument can be understood as a subset of persuasion. For Aristotle, rhetoric (written somewhere between 360 and 334 BC) is strictly equivalent to the study of the modes of persuasion, which involves three major components in communication: the speaker, the audience, and the content of argument. In order to make a persuasive speech, a persuader must take into account three points: the means or sources of persuasion, the style or language, and the arrangement of the speech. The interaction of the modes of persuasion means that audience can be persuaded by the personal character of speaker (ethos), by working on their emotions (pathos), or by the proof provided by the speech itself (logos). Based on this systematic account of rhetoric, argument is, therefore, a means to persuade using logic (logos) or reasons or verifiable proof.

Persuasion is then a broad term and it may not be easy to distinguish it from argumentation. For ease of understanding and learning purpose, a number of composition textbook authors and communication professors have attempted to distinguish argumentation from persuasion. Nadell, Langan, and McMeniman (1993), for example, define argumentative writing as a piece of writing that uses "logic (logos) to convince readers of the soundness of a particular opinion on a controversial issue" (p. 571). Persuasion, on the other hand, uses "emotional language and dramatic appeals (pathos) to readers' concerns, beliefs, and values" (op. cit.). Several other authors' definitions and explanations also fall within such categorization (See, for

example, Larson, 1988; Langan, 1996; and Smalley, Ruetten, and Kozyrev, 2001). The major characteristics that often distinguish persuasion from argumentation, as these writers have pointed out, are that persuasion not only encourages readers to accept an opinion but also often influences them to commit themselves to a course of action while the latter focuses on the use of logic to support the writer's position on a controversial issue. Historically, however, the major motive of all these attempts to define persuasion and argumentation and to distinguish them might have been influenced by the dropping of the teaching of the two modes of persuasion, the appeal to writer's credibility (ethos) and the affective appeal (pathos) after since the writing curriculum of England and the United States substituted persuasion with argument in the eighteenth century (Connor, 1996).

Indeed, in a persuasion-filled world, people usually yield to two forms of persuasion, logical appeals and emotional appeals, considering mass media or television commercials. By realizing that argumentation and persuasion are usually blended, writers can make their arguments compelling by employing a persuasive strategy using emotional appeals and positively charged language. Nadell, Langan, and McMeniman (1993) suggest that when writers prepare argumentation-persuasion essays that advance the writers' position through a balanced appeal to reason and emotion, emotion supports rather than replaces logic and sound reasoning. Cicero also recommends this as early as in the period of ancient Rome. In his model of argument, he demands that the propositions (writers' claims or positions) be proved rationally though he reserves the concluding as the most climatic portion for emotional appeals (cited in Nicholas and Nicholl, 1994).

In this research, based on the analysis of collected data (students' argumentative essays), it was most likely that student writers resorted more to rational appeals (empirical scientific proof) than to emotional appeals though not at the exclusion of the latter. This might be due to two things. First, students were generally encouraged to write on the topics or the subject matters in which they were specialized according to their fields of study at the university; and second, the proportion of a persuasive strategy, logical reasoning and affective appeals adopted in their essays might be varying depending on the complexity of individually selected issues and forwarded propositions. For instance, the claim proposed by one student with a major in public health that the new Medicine Act enacted by the Ministry of Public Health of Thailand should be passed by the government as a legislation for controlling the prescription of drugs by physicians and pharmacists which would benefit patients in a number of ways might require the writer to play on the target audience's feelings, concerns, or beliefs than to resort more to scientific proof to support her proposition. A persuasive quality of this student writing depends on how well and accurately she could base her argument on the common ground existing between she as the persuader and the target audience, making certain assumptions about the audience and its beliefs (pathos) and making use of such projected assumptions in stirring audience's emotions through the use of emotionally charged language.

2.1.2 Structure of Argument

"Mere knowledge of truth does not give you (the writer) the arts of persuasion," said Plato in *Phaedrus* (parenthesis added). In order to argue effectively and persuasively, Aristotle wrote in *Rhetorica*, the speaker or writer has to pay

attention to three things: the means of persuasion (ethos, logos, and pathos), the language or style to be used, and the proper arrangement of the various parts of the argument. This implies that simply having opinions and supporting facts are not enough for effective argument. The writer must also decide how to organize and express them and how to counter opponent's objections using appropriate tone, voice, and language for the audience (Vesterman, 2000).

2.1.2.1 What Classical Rhetoric Says

Given that the key principles of argumentation-persuasion adopted and applied in today's English composition instruction and textbooks can trace its roots to Aristotle's text and classical rhetorical theories (Connor, 1996), it is worthwhile to discuss briefly what classical rhetoric says about the structure of argument. In Aristotle's Rhetoric, a speech contains two parts: the statement of the case, and the argument. The essential features of a well-arranged speech cannot in any case have more than Introduction, Statement, Argument, and Epilogue. Refutation of the Opponent is part of the arguments (1414a, 1414b)[1]. After Aristotle, Cicero (106-43 BC), influenced by his study of persuasion (Rhetoric) provided another one of the oldest systematic layout of argument though his version was more elaborated aiming at its written form. For Cicero, an argument usually contains six elements: 1) Introduction; 2) Background; 3) Partition, or statement of propositions; 4) Confirmation proof of propositions; 5) Refutation; and 6) Conclusion, or appeal to sympathy (Nicholas and Nicholl, 1994).

---

[1] The pages and columns of the standard Berlin Greek text as assigned in the Oxford translation (Adler, 1990).

2.1.2.2 Argument Strategies

While there are a variety of techniques and style for argumentative writing, the strategies reviewed and discussed in this section are crucial for basic university writers particularly for EFL students. Drawn vastly from several writers, the following summarizes some effective techniques suggested to be useful for students learning to write arguments.

- **Context Building**

Since an argument does not occur in a vacuum, the introductory part of an essay has an important function in setting the context for the argument. In order to develop an effective argument or introduction in particular, the writer must analyze and understand the rhetorical situation (Bitzer, 1968) or the elements that combine to constitute a communication situation. Based on Wood (1998), a rhetorical situation of any argument has five elements:

I.   **Exigence:** The real life situation that happens to cause the argument; a problem to be solved; a situation that requires some modifying response from an audience.

II.   **Audience:** The targeted audience who cares enough to listen, read, and pay attention, and to change its perceptions as a result of the argument.

III. **Constraints:** The existing circumstances, beliefs, attitudes, values, or traditions that limit or influence the writer to write in certain way or that cause the targeted audience to respond to the situation in a particular way.

IV. **Author:** The writer who writes an argument in response to the exigence of the situation; the writer's background, experience, education, and values.

V. **The text:** The unique characteristics of the written argument such as the format, organization, argumentative strategies, and language.

Taking into account the specific rhetorical situation of any argument, the introductory part should orient the readers to the argument by providing them with sufficient information to understand the issue being argued. According to Batteiger (1994) and Clark (1994), this can be done in a variety of ways by including in the introduction the following:

- Demonstrate that a problem exists that needs solving

- Provide a background summary of the topic, defining relevant terms

- Develop a scenario that presents a variety of solutions that will solve the problem, including the writer's proposed solution

- Indicate why the topic is important or why a particular solution is, or would be, beneficial or preferable to the alternative solutions

- State thesis or claim and map out ongoing arguments

- **Reasoning**

In argumentative academic writing, writer can support their arguments using either appeals to reasons or appeals to emotion. Since reasoning is integral to any argument, using appropriate and reliable evidence to support arguments will convince the readers as well as contribute to the writer's credibility.

For the topics or issues related to physical facts such as engineering and technology subjects, the writer usually relies mainly on inductive reasoning strategy (scientific method) using verifiable reasons or demonstrative proof to support his or her arguments or generalizations. Such scientific proof involves facts, examples, statistics, and expert authority taken from books, articles, reports, interviews,

documentaries, as well as electronic sources such as the Internet. Reid (1989) suggests that sound evidence should be recent, specific, unified, sufficient, and representative since insufficient or atypical evidence often leads to hasty generalizations that mar the essay's logic, weakening rather than holding the writer's position.

- **Acknowledging Differing Viewpoints**

In addition to using convincing evidence to prove the writer's claim, the other way that the writer can establish his or her credibility (ethos) in argumentative writing is to represent and evaluate the opposing points of view on the issue fairly and accurately. There can be three ways that the writer should acknowledge opposing arguments (Langan, 1994). One basic but effective technique is to cite the opposing viewpoint in the thesis statement using a two-part proposition consisting of a subordinate clause followed by a main clause. The first part of the proposition (the subordinate clause) acknowledges the other side's viewpoint; the second part (the main clause) states the proposition. For instance, "Although some computer programmers may argue that JAVA language has a problem in its compiling process, novice as well as experienced programmers should use JAVA in programming web-page contents because of its efficiency, high security, and versatility."

A second technique is to use a paragraph within the body of an essay to summarize and refute the opposing opinions in greater detail (See Figure 2.1). A refutation section can take two forms (Nadell, Langan, and McMeniman, 1993). The writer can choose either to mention all the opposing viewpoints and then present

counterargument to each of those points or to present and refute one opposing point at a time.

2.1.3 Organizational Plans for Argumentative Essays

Statements of argument can be formal or informal in design. An argument column in a specialized published material, says 'Opinion and Analysis' section in the Bangkok Post, is less likely to have a fixed structure. For argumentative writing in academic context, an argument is likely to be tightly organized; thus, an organizational plan is desirable. Figure 2.1 illustrates three basic organizational plans for argumentative essays found in most English composition textbooks (See, for instance, Reid, 1988; Smalley, Ruetten, and Kozyrev, 2001).

Figure 2.1 Three Basic Organizational Plans for Argumentative Essays

Plan A

I. Introduction (+ thesis statement of intent)
II. Background paragraph about topic (Optional: depending on assignment, audience, and the available material)
III. Pro argument #1 (weakest argument that supports the opinion)
IV. Pro argument #2 (stronger argument that supports the opinion)
V. Pro argument #3 (strongest argument that supports the opinion)
VI. Con (Counterarguments and refutation)
VII. Solution to the problem (Optional: depends on assignment, audience, and the available material)
VIII. Conclusion (summary + solution, recommendation, or call to action)

Plan B

I. Introduction (+ thesis statement of intent)
II. Background paragraph about topic (Optional: depending on assignment, audience, and the available material)
III. Con (Counterarguments and refutation)
IV. Pro argument #1 (weakest argument that supports the opinion)
V. Pro argument #2 (stronger argument that supports the opinion)
VI. Pro argument #3 (strongest argument that supports the opinion)
VII. Solution to the problem (Optional: depends on assignment, audience, and available material)
VIII. Conclusion (summary + solution, recommendation, or call to action)

Plan C

I.   Introduction (+ thesis statement of intent)
II. Background paragraph about topic (Optional: depending on assignment, audience, and available material)
III. Counterargument #1 + Pro argument to refute it
IV. Counterargument #2 + Pro argument to refute it
V.   Counterargument #3 + Pro argument to refute it
VI. Solution to the problem (Optional: depends on assignment, audience, and available material)
VII. Conclusion (summary + solution, recommendation, or call to action)

Source: Reid (1988), p. 94

Reid (1988) suggests that in the assignment of approximately 700 to 900 words, the essay will have four to six paragraphs. In the figure, in each plan, one or more of the paragraphs can be optional; that is, the student writer will need to choose either to use or not to use that paragraph depending on the required length of the assignment or the complexity of the arguing issue. She further notes that for most short argumentative essays, the introductory elements (introduction, narration, and partition) all are usually combined in one introductory paragraph including a thesis statement.

Based on the basis that the elements of arguments can be organized in various ways, Webb (1994) provides a guideline of the four ways for organizing written arguments, depending on the situation and the writer's purpose. One commonly used classic pattern begins by stating the thesis in the introduction, continues to present the evidence to support the thesis, next addresses and refutes any likely objections, and concludes with a restatement of the thesis with a call to action. In the situation where the writer is arguing in response to someone else's argument, it is suggested that instead of beginning with a statement of thesis, the essay may be best to begin by summarizing the opponent's position followed by the writer's thesis and

argument. Yet another pattern is for arguing in favor of a policy or course of action that the writer believes will solve a problem. The essay may need to begin by showing that there is a problem before the writer goes on to present and defend a proposed solution. The other final plan provided by Webb is for writing for an apathetic, skeptical, or hostile audience which is obviously difficult to convince since this type of reader "might find the thesis distasteful enough to dismiss it without giving a writer a chance to defend it" (p. 340). In this situation, it is better off to delay a statement of thesis until the end of an essay, after evidence is presented. For a hostile audience, Nadell, Langan, and McMeniman (1993) cautions the writer to resort mainly to logical reasoning or "hard-to-dispute facts" and avoid using "emotional appeals which might seem irrational, sentimental, or even comical" (p. 577).

Webb's guideline of organizing plans is not exhaustive. The other important kind is the so-called Rogerian argument[2] (Young, Becker, and Pike, 1970) which is suggested to be more effective than traditional argument when the issue is particularly highly sensitive or controversial and the audience the writer tries to persuade is, in fact, the opposition or the writer's adversaries. Rogerian argument, as opposed to traditional argument that intends to win the argument, aims to reconcile conflict of interests, to achieve a mutual communication, to work toward changes in both sides as a means for achieving common ground, and to reach a compromise position or solution. Because of these extremely important features of Rogerian strategy, it is suggested that instead of following set organizational patterns as

---

[2] Named after Carl Rogers, a psychotherapist who invented the empathetic listening technique and used it in psychological counseling and in improving communication in difficult, emotionally charged situations. Young, Becker, & Pike (1970) applied Rogers' ideas to formulate Rogerian argument, "a method for helping people in difficult situations to make connections, create common ground, and understand one another" (cited in Wood, 1998, p. 249).

illustrated above, the writer may need to begin the essay by introducing the issue and

showing the opponent's position is understood, next showing in which contexts and

under what conditions the opponent's position is valid, then stating the writer's

position includingthe contexts in which it is valid, and concluding by showing how the

opponent's position would benefit if the opponent were to adopt the writer's position

(Wood, 1998). Table 2.1 summarizes what has been discussed so far and contrasts

Rogerian argument with the traditional pro-and-con model of argument.

Table 2.1 Traditional and Rogerian Argument Compared

	Traditional Argument	**Rogerian Argument**
Basic Strategy	Writer states the claim and gives reasons to prove it. Writer refutes the opponent by showing what is wrong or invalid.	Writer states the opponent's claim and points out what is sound about the reasons used to prove it.
Ethos	Writer builds own character (ethos) by citing past experience and expertise.	Writer builds opponent's character perhaps at expense of his or her own.
Logos	Writer uses logic (all the proofs) as tools for presenting a case and refuting the opponent's case.	Writer proceeds in an explanatory fashion to analyze the conditions under which the position of either side is valid.
Pathos	Writer uses emotional language to strengthen claim.	Writer uses descriptive, dispassionate, neutral language to cool emotions on both sides
Goal	Writer tries to change opponent's mind and thereby win the argument.	Writer creates cooperation, the possibility that both sides might change, and a mutually advantageous outcome
Use of Argumentative Techniques	Writer draws on the conventional structures and techniques	Writer throws out conventional structures and techniques because they may be threatening. Writer focuses, instead, on connecting empathetically.

Source: Reproduced from Wood (1998), p. 250

2.2 EFL Writing in an Integrated Process-Genre Approach

The discussion on genre teaching has been unfortunately polarized as reflected in a number of publications in the field of teaching L1/L2 writing. Two lines of debates exist in the teaching of genres of writing; one focuses on whether genre should be explicitly taught or learners acquire genres tacitly, and the other on the efficacy of the so-called genre-based approach advocated by ESP/EAP professionals and those who work within the Sydney School tradition. A substantive debate on whether teachers should teach genre explicitly can be found in the October 1993 issue of Research in the Teaching of English (See the contributions by Freedman, Williams and Colomb, and Fahnestock). For the discussion on the hypothetical effectiveness of genre-based approach to teaching writing, the recent special edition of Journal of Second Language Writing (See volume 12, 2003) provides a discussion forum on the current best thinking as to the adequacy of both process and genre approaches to teaching written genres. These conflicts of interest in methodology issues, as some critical language educators have envisaged, are generating more heat than light as they are based on narrowly- conceived views of how each approach conceptualizes writing and learning to write. Tribble (1996) sees that writing generates "many, often conflicting, views" (p. 37). In Badger and White's (2000) perspectives, the conflict between the various approaches is misguided and damaging to classroom practice. They further note that the approaches are indeed largely complementary considering each approach's pedagogical merits and how each of them (product, process, and genre approaches) can contribute to students' development of writing ability.

In order to resolve the conflicting views as well as to respond to a repeatedly articulated dissatisfaction with the limitations of the concept of *method* in the 1990s,

Kumaravadivelu (2001) along with other colleagues (e.g., Richards, 2001; Clarke, 1994; Prabhu, 1990) proposes that the L2 profession is faced with an imperative need to construct what he refers to as a *postmethod* pedagogy. Central to the concept is the need to go beyond the limitations of the concept of method which is often conceived of as one exclusive single construct guiding the form and function of every component of L2 pedagogy including curriculum design, syllabus specifications, materials preparation, instructional strategies, and testing techniques. Postmethod pedagogy, Kumaravadivelu argues, like all politics, is local that it has to take into account local exigencies or a holistic interpretation of particular situations including a particular group of learners pursuing a particular set of goals within a particular institutional context. In turn, for practicing teachers, this means reflective teaching that involves observing their teaching acts, evaluating their outcomes, identifying problems, and finding solutions. In other words, the postmethod pedagogy is the localized pedagogy that suggests the need to draw from a variety of methods to be used to meet the needs of learners in particular exigencies. Only by taking up this view, the conflicts of interest in the ideologies of what to teach and how to teach it between those who support the latest approach and those who reject it then can be reconciled.

Along the notion of postmethod pedagogy came what is referred to recently in the field of teaching of writing as *post-process*. This is evident in a series of discussion on the issue "L2 Writing in a Post-Process Era" in a recent special edition of Journal of Second Language Writing in 2003 (See volume 12, 2003). While the focus of the preceding debate on whether teachers should be explicit or implicit in the teaching of written genres tends to represent a standoff between those who advocate explicit teaching and those who believe that learners learn genre subliminally, the

discussion on the post-process issue does not necessarily mean a call for a paradigm shift again like what happened in the past when there was a call for product-driven syllabus to be replaced by process-oriented writing instruction approach. Rather, suggestions are made on the necessity and legitimacy of process and genre approaches in teaching students curriculum genre, and this can be seen as representing a more compromised circumstance in which both approaches are interdependent rather than mutually exclusive. Insofar as the post-method post-process perspective is the current concern in the teaching of L2 writing, it might be more fruitful to discuss L2 writing pedagogy with a backdrop of local exigencies than in a way that represents rigid dichotomies.

According to Tribble (1996), for the writer to write effectively, he or she has to have content knowledge, context knowledge, language system knowledge, and writing process knowledge. In other words, it can be argued that if writers know what to write in a given context, what the reader expects the text to look like in a given context, and which parts of the language systems are relevant to the particular task in hand, and have a command of writing skills appropriate to this task, then they have a good chance of writing something that will be effective. Since it is hardly for any single approach to teaching writing to address all the types of knowledge the writer needs to know, a range of knowledge that writers need to know to produce an effective prose text might suggest ways in which a combination of approaches to writing instruction can be beneficial in enabling students to write in a given genre. Research has consistently indicated that while nonnative writers often find it difficult starting to write, they also bring to the writing tasks both linguistic and rhetorical deficits (See Section 1.2). An effective methodology for writing then needs to incorporate the

insights of process and genre approaches to help students overcome their writing difficulties (e.g., Badger and White, 2000; Flowerdew, 2000; Kay and Dudley-Evans, 1998; Coe, 1994). Badger and White (2000) suggest that one way of doing this is to start with one approach and adapt it. In practice, there is no doubt that the process approach is useful in helping students develop viable strategies for getting started (finding topics, generating ideas and information, focusing, and planning structure and procedure), for drafting (encouraging multiple drafts), for revising (adding, deleting, modifying, and rearranging ideas), and for editing (attending to vocabulary, sentence structures, grammar and mechanics).

While writing in a process approach would provide students with knowledge of context and skills of using language, the case is also made for students to be taught explicitly knowledge of rhetorical forms and conventions of the genre of writing task. ESL students, even more than native English-speaking students, must be taught "the styles of thinking and ordering that dominate U.S. academic discourse" (Shaughnessy, 1977, p.239). Focusing predominantly on the process rather than on the qualities required for the product puts students at a disadvantage outside ESL or English composition programs because evaluation in the larger academic environment remains focused on the end product (Leki and Carson, 1997). More recently, Chang and Swales (1999) argue that academically oriented non-native speakers (NNSs) need to be made aware of the rhetorical and sentence-level features prevalent in formal written discourse. They suggest that even for advanced and highly literate NNSs, exposure to substantial amounts of reading and experience with writing in the relevant context does not ensure an awareness of the necessary discourse and sentence-level linguistic features of writing. The authors' teaching implication is that explicit instruction in

advanced academic writing is needed. "It remains highly unlikely that the unmodified process approach to teaching composition will prove to be sufficient" (Grabe and Kaplan, 1996, p. 173).

Hyland (2003) raises one instance among others in the process model that might disempower teachers and cast them in the role of well-meaning bystanders (see also Cope and Kalantzis, 1993). Given that the role of teachers in the approach is reduced to his or her involvement in developing students' metacognitive awareness of their writing processes and responding to writing, students are offered no way of seeing how texts are codified in distinct and recognizable ways in terms of their purpose, audience, and message because language and rhetorical organization tend to be things tacked on to the end of the process as editing (Macken-Horarik, 2002). "In the process writing conference, for example, the student acts and the teacher reacts; the writer follows language, and the teacher follows the writer (Murray, 1982, p. 163, cited in Kalantzis and Cope, 1993). "Everything is reversed. I have to give up the active, non-delegating, pushing, informing role for another kind of activity, the activity of waiting. Action in conferences is redefined as intelligent reaction. The child must lead, the teacher intelligently reacts" (Graves, 1983, p. 127, quoted in Kalantzis and Cope, 1993).

By taking into account that L2 writers need to be familiarized to the discourse features of genre, the genre-based approach then seems to have an important place in the teaching of writing. However, work on genre analysis and its applications in the teaching of writing can be divided into two groups. The first group consisting of those who work in Swalesian (ESP) tradition (e.g., Swales, 1990; Dudley-Evans, 1987; Bhatia, 1993) tends to focus pedagogically on the tertiary level and beyond in enabling

students to produce the genres required in their professional and academic study such as business and job application letters, research articles, and laboratory report. The other group works in the Australian tradition (often referred to in North American and British publications as the Sydney School) in the teaching of genre of spoken and written language (e.g., Cope and Kalantzis, 1993; Hammond, Burns, Joyce, Brosnan, and Gerot, 1992; Feez and Joyce, 1998). In contrast to the ESP tradition, the genre-based approach as practiced in Australian context tends to be more explicit in its pedagogical framework as reflected in the use of 'curriculum genre' in institutional settings across different language proficiency levels (Mary Macken-Horarik, 2001).

Genre-based pedagogy is often identified in the form of a teaching and learning cycle first developed in Australia through the work of educational linguists and educators who have been working with disadvantaged groups of students (Figure 2.2). The cycle consists of three stages: context building and modeling, joint text-construction, independent construction of text. To teach genre, the teacher and students work through a series of stages and classroom activities in the cycle in order to learn the target genre. Firstly, the genre is introduced through a model text that exemplifies the genre. The emphasis at this stage is on the text's social purpose (functions), how the information in the genre is organized (schematic structure), and aspects of the way the text speaks (lexico-grammatical features). Secondly, a text of the genre is constructed jointly the teacher and students. The teacher acts as a scribe as the students contribute to a jointly constructed text, which approximates the schematic structure of the genre and employs its key lexico-grammatical features. In the third and

final phase, students choose and research a topic, and prepare to construct their own texts of the genre; drafting and conferencing with peers and the teacher for feedback on their individual writing efforts. In theory, the cycle is intended to be so flexible that it is possible at any time to return to activities from earlier stages of the cycle if students need revision or further practice in order to progress (Callaghan, Knapp, and Noble, 1993). In short, an integral aspect of the genre approach is working with the whole texts at the beginning, requiring that before attempting to write in the target genre, the students need to be exposed to the genre by reading, analyzing, and discussing examples of it. The method offers writers an explicit understanding of how and why texts in the target genres are structured and organized in certain ways to achieve their communicative social purposes (see Cope and Kalantzis, 1993 for a detailed account on the cycle).

Figure 2.2 **Stages of the Teaching and Learning Cycle**
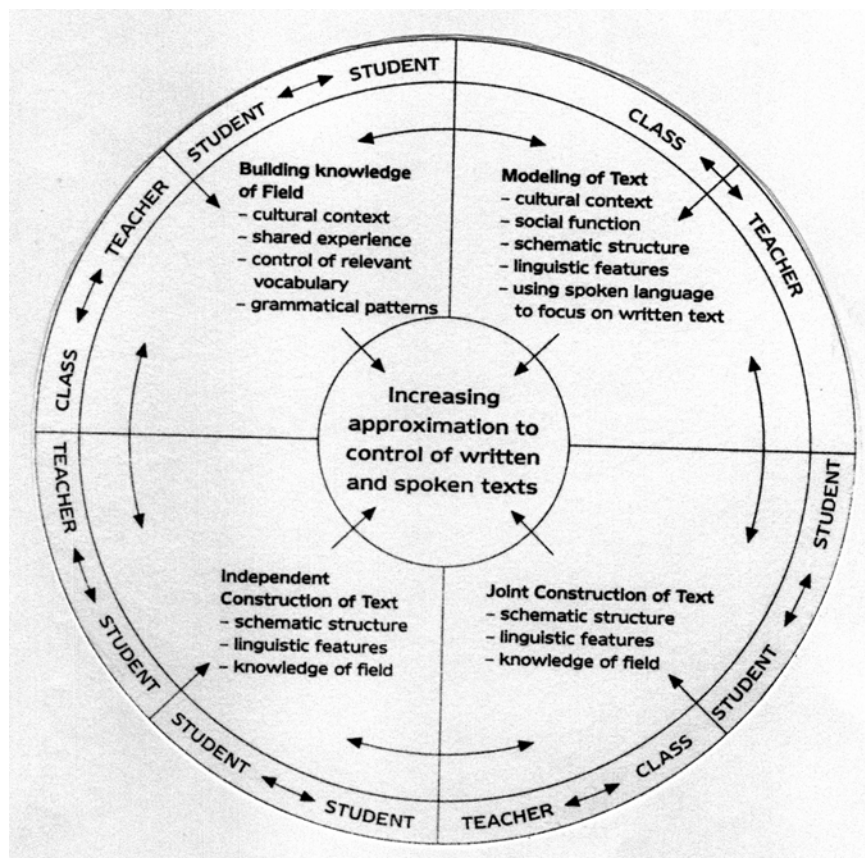
The theoretical underpinning of the genre approach rests on the ideas that 1) learning occurs more effectively if teachers are explicit about what is expected of students and 2) the process of learning language is a series of scaffolded developmental steps (Feez and Joyce, 1998). As opposed to the invisible pedagogy of process approaches (see also critique by Hyland, 2003 and Macken-Horarik, 2002), genre pedagogy represents a visible pedagogy (Bernstein, 1990) in which what is to be learned and assessed is made clear to students. The design of the teaching-learning cycle draws on the work of Vygotsky (1978) and its interpretation by Bruner (1986). Vygotsky, as a constructivist, proposes that knowledge itself is structured and developed within the individual through active learning. Through interaction, the individual (learner) progresses from what Vygotsky calls an actual developmental level to a potential developmental level. Between these two levels is the zone of proximal development or "the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance of and in a collaboration with more capable peers" (1978, p. 86, cited in Richard-Amato, 2003).

A model of learning based on Vygotsky's insights suggests that through a dynamic social process, the teacher in a dialogue with students can focus on emerging skills and abilities, engage with learners in challenging significant tasks, observe what students can do independently, and give them help and guidance in moving forward to solve problems, thereby forming a type of cognitive apprenticeship (Wells, 1999; Rogoff, 1990). Bruner (1986) uses the term scaffolding to describe the teacher's role

in the interactive collaboration. As reflected in the cycle, novice L2 writers seem to require greater support during the early stages of working with a new, unfamiliar genre and less later. The teacher takes an authoritative role by intervening in the initial stages through providing information, appropriate language input, and opportunities for guided practice. As learners achieve greater control of the genre, support from teacher is gradually removed and more responsibility shifted to learners in constructing the genre themselves (Burns, Joyce, and Gollin, 1996).

Gray (1987, in Feez, 2001) lists four points that teachers need to consider as they implement genre pedagogy:

1. Joint construction involves negotiation between the teacher and the learner, not domination by the teacher. Both the teacher and the learner need to have a shared understanding of the context and of the meaning being negotiated. Problems are shared, and the teacher only asks direct questions when learners show they have a chance of success. Throughout this process, learners are thinking and making choices to contribute to the joint construction.

2. Teachers have to create contexts in which the use of the target language is legitimate and meaningful.

3. Jointly constructed and negotiated meanings are best supported or scaffolded within predictable and familiar routines of interaction and activity over extended periods of time.

4. Teachers use scaffolding to monitor the level of difficulty as control is gradually handed over to the learner.

By utilizing what is learned about writing and learning to write, an integrated process-genre approach can be a promising alternative that reflects the current best thinking in the field of L2 writing with respect to pedagogy. As an endnote to this section, what Widdowson (1990) says might provide a strong basis for the legitimacy of genre as a visible, explicit pedagogy:

> The whole point of pedagogy is that it is a way of short-circuiting the slow process of natural discovery and can make arrangements for learning to happen more easily and more efficiently than it does in natural surroundings… Pedagogy is bound to be a contrivance; that is precisely its purpose. If what went on in the classrooms exactly replicated the conditions of the world outside, there would be no point in pedagogy at all. And … the advantage of pedagogy is denied if it just leaves learners to learn by doing without quite deliberately contriving ways of assisting them in getting to know the language system at the same time, as the essential resource for their doings (p. 162, in Feez and Joyce, 1998).

2.2.1 Some Empirical Research on Explicit Teaching of Genre

Reznitskaya (2002) reports the positive influence of collaborative discussions (collaborative reasoning) and explicit instruction on the acquisition and transfer of argumentative knowledge. As a quasi-experiment, fourth graders as the experimental group and fifth graders as the controls completed the same argument-related tasks, after receiving different instructional treatments. The tasks included an interview designed to elicit students' awareness of argumentation criteria, a persuasive composition, and a recall of an argumentative text. The findings reveal positive gains in learning and transfer as shown by the students' developed competence in performing the argumentative tasks. Students who engaged in discussions with or without explicit instruction provided well-articulated responses to the interview questions, showing a rather sophisticated understanding of argumentation functions and criteria. Overall student performance on the persuasive essay was positively affected only by participation in discussions. The recall of an argumentative text was

insensitive to variations in treatment. While the results of the experiment were mixed, the author emphasized that even when overall effects were not present, the oral and written productions of some students especially of diverse ethnic groups suggested benefits from discussions and explicit instruction in argumentation.

Yeh (1998b) reports similarly the effectiveness of explicit instruction in the requirements for particular thesis-support pattern of argumentation. The study was based on a hypothetical perspective on learning that many students particularly those cultural minority middle-school students participated in the study benefit from explicit instruction in heuristics regarding argument structure, in addition to exposure and immersion in contexts where the particular pattern is appropriate. Using pretest-posttest control group design, the study contrasted explicit instruction in heuristics for constructing arguments in combination with immersion in debate and peer response activities with a version of the same approach that excluded explicit instruction and included only immersion activities. The resulting outcomes of the study showed that students in the experimental group applied the heuristics flexibly, implying that they learned principles than rote procedures for argumentation and were able to adapt the heuristics and transfer their knowledge to a range of topics. The experimental group students also demonstrated greater knowledge of argument criteria and strategies regarding students' judgment on the content and organization needed to generate logically connected arguments. In a broader sense, the implication of the study suggests that clarifying the requirements for writing argumentative essays such as the heuristics tested in the study can enhance traditionally underprepared students' ability to write academic essays.

A study carried out at the National University of Singapore reports positive results of a genre-based academic writing course in which a class of second-year English major undergraduates were trained in the structural and interpersonal aspects of argumentation (Varghese and Abraham, 1998). The structural aspects were taught and measured using Toulmin's (1958) framework for argument analysis including the quality of claims, grounds, and warrants used. The interpersonal aspects of students' essays were also assessed in terms of the creation of clear persona, audience adaptiveness (the appropriate use of rational and emotional appeals), and stance towards the unique discourse of argumentation. With the one group pretest-posttest design, the study required the students to write two argumentative essays before and after exposure to explicit instruction in argumentation. The findings revealed pre- to posttest gains in students' abilities to formulate claims, to offer specific and developed grounds, and to use more reliable warrants. Students' essays also showed improved and more effective use of interpersonal aspects of argument, building better writer credibility, developing fuller rational and emotional appeals, and conveying both sides of an argument in order to resolve the problem.

One study reports on a case study of fifth-grade ESL learners taking a five-week social studies course that focused on narratives, descriptions, expositions, and persuasive texts (Reppen, 1995). Students were introduced to each genre through the content material and different writing tasks and the instruction that combined the writing process approach, language arts skills activities with specific content material, and direct instruction on different genre forms. Pre- and post- assessment measures used to assess change in student writing, content knowledge, and attitudes reflected a positive change. With respect to the persuasive task, it was found that students

realized the need to view the task from a perspective other than their own, to anticipate an argument and generalize reasons that would support their stand. In Reppen's perspective, a genre-based approach to content writing instruction can offer ESL students useful practice in school-valued ways of writing while also learning content material and working through steps in the writing process. She also suggests that a focus on the genre demands of different ways to organize information helps increase student awareness of how different ways of organizing information in writing interacts with the purpose of the text and this is an important step in helping students become more successful writers.

A posttest-only control group study carried out at the Tokyo Sangyo University in Japan reveals positive learning gains as a result of formal explicit instruction in rhetorical patterns of argumentative writing in addition to the teaching of the basic elements of argumentation such as audience expectations (Yoshimura, 2002). Student essays were analyzed and compared at the end of the course. The largest significant gains were found in the experimental group in discourse-level fluency and rhetorical proficiency as measured by three judges using Jacob's (1981) analytic rating scales and t-unit analysis.

2.3 Assessing Argumentative Writing

The key issues in L2 writing assessment and research in the field of language testing can be viewed to involve a consideration of the following aspects in any context of testing writing: the test taker (writer), the task or prompt, the written text, the rater(s), and the rating scale (Hamp-Lyons, 1990; Weigle, 2002). A complex network of effects as a result of an interaction among these components in a language

assessment situation from which the test taker's score derives has been so far beyond practitioners' effort to control (Hamp-Lyons, 1990).

In rating based assessment such as the judging of students' written ability, concern has been overwhelmingly with the marking process or the estimation of reliability including the consistencies of scores assigned by different raters on the same essays (inter-rater reliability) or by the same raters on the same essays on different scoring occasions (intra-rater reliability) rather than with the reliability of the essay test (task/prompt) itself (reliability of essay test) or the variation between the writing tasks or within students (Pollitt, 1991; Hopkins, 1998).

Be this as it may, Alderson (1991) suggests that in rating based assessment, it little matters or even does not matter whether test examiners can control all the aspects of testing especially with respect to validity and reliability as concerned by item based tests (indirect tests) if the theoretical advantage of rating based assessment is to be realized. He stresses instead that what matters very often is whether the score that is reported reflects the test taker's ability no matter whether the score is derived by rating or by counting. In assessing performance tasks such as writing, McNamara (1996) points out in addition that the scale usually reflects the scale developer's notion of what skills or abilities are being measured by the writing task. Based on this reason, the development of a scale and its descriptors are of critical importance for the validity of the assessment.

While a score in a writing assessment is the outcome of an interaction that involves the writer, the task, the rating scale, and the rater(s), two of these elements are of central importance in scoring: defining the rating scales and ensuring that raters use the scale appropriately and consistently (Weigle, 2002). The issues concerning the

measures used for ensuring rater reliability and for establishing reliability of the ratings are treated in some detail in the next chapter. Readers are referred to Hughes (1989), Cohen (1994), or Hyland (2003) for a good overview of issues in assessing L2 writing, and to Weigle (2002) for practical advice for designing tasks and scoring procedures for writing tests. This section turns to focus on the development of rating scales and scale descriptors, and relevant empirical studies with a particular emphasis on assessing student argumentative writing.

2.3.1 Writing Scale Descriptors

Weigle (2002) indicates two main approaches to writing scale descriptors. One approach advocated by Bachman and Palmer (1996) is that the scale descriptors can be done a priori or by defining in advance the ability being measured and then describing a number of levels of attainment from none to complete mastery. The advantage of this approach is that each score level's description is useful in reporting what a test taker is capable of doing with the language, allowing one to make inferences about a test taker's language ability on an absolute scale rather than relative to other test takers. However, a potential drawback of such absolute scale, according to Bachman and Palmer (op.cit.) is that its descriptions tend to represent imprecise distinctions between the levels (e.g. words that indicate levels of ability/mastery such as 'none,' 'limited,' 'moderate,' and so on). In addition, it is likely that inexperienced or novice raters may have difficulties understanding and making distinctions reliably without extensive training and repeated exposure to texts that instantiate the various scale levels.

Another approach is to generate scale descriptions empirically through the examination of actual scripts and operational ratings of writing performance. This pragmatic approach advocated by North and Schneider (1998) suggests five data-based methods of scale construction that involve expert judgements of the key features at different levels of performance, statistical analysis of ratings in relation to scale descriptors, or textual features of performances at different levels. The most common process of scale construction using this data-based method involves gathering sample scripts on a writing task from students at all relevant levels of proficiency and, with a group of instructors familiar with the proficiency levels, defining the characteristics that differentiate the samples. In this approach, the definitions of scale levels take the form of verbal description rather than levels of mastery. The verbal descriptions often refer to notions that do not lend themselves to the levels-of-mastery approach such as audience awareness, ideas and arguments, and overall communicative effectiveness. Two typical examples of rating scales developed under this empirical method are the TOEFL writing scoring guide and Michigan writing assessment scoring guide (See Weigle, 2002 for a variety of types of scoring guides).

Given the choice between these two approaches to constructing scales, which one to choose depends on the theoretical standpoint adopted by test administers whether they see that the most important aspects of the ability being tested can be measured on a scale of none to complete mastery or on factors related to the purpose of the assessment. It is suggested that the mastery approach as advocated by Bachman and Palmer (1996) is particularly appropriate if the purpose of writing assessment is to make inferences about an inherent ability (i.e. a student *has* ability X) rather than in terms of a pragmatic ascription (i.e. a student *can* do X, without reference to the exact

nature of the underlying ability). However, the pragmatic approach as proposed by North and Schneider may be useful when the outcome of assessment is to predict how someone will perform on future similar tasks rather than other task types.

2.3.2 The Rhetoric/Syntax Split

Kroll (1990a) calls practitioners to put into consideration in evaluating ESL student writing the two notions, rhetoric and syntax, claiming that ESL writing often has a split between accuracy of language use and fluency of ideas. Traditionally, accuracy can be defined as the focus on discrete elements of the rules of language, while fluency has been referred to as the focus on the communication of ideas without consideration of discrete language elements (Lennon, 1991b, in Reid, 1993).

Kroll's analysis of the accuracy versus fluency debate in language teaching demonstrates that ESL writing proficiency can be broadly divided into "plus-syntax and minus-syntax," and "plus-rhetoric and minus-rhetoric." She defines syntax as the facility to use the grammatical system of standard edited English in such categories as sentence structure, word form, word order, verb tense, etc. In contrast, a piece of writing with rhetorical competence 1) limits and focuses on the topic in a manner appropriate to its overall approach and length, 2) remains focused on the topic throughout, 3) creates and uses paragraphs effectively, 4) maintains a consistent point of view, 5) sequence ideas in a logical manner, and 6) uses coherence and cohesion devices appropriately and as necessary (1990a, p. 43).

In reading student writing, according to Kroll, "one paper can provide insightful commentary on a substantive topic while replete with problems in spelling and punctuation [+ rhetoric and – syntax]. Another paper can exhibit a wide range of sentence structures, flawless syntax, adherence to mechanics, yet lack development

and support of its central thesis [- rhetoric and + syntax]" (p. 40). Still another paper may be hard to read because it contains both second language errors and a lack of coherence [- rhetoric and – syntax].

Kroll's work is based on an investigation of an essential difference between international students who study in the United States and non-native speakers of English who become residents in the U.S. The important difference between these two groups of learners is that students who have studied English as a foreign language (EFL), and who have entered post-secondary institutions in the U.S. following several years of EFL study in their native countries ordinarily have acquired English *visually,* while those immigrants who have lived and attended elementary and/or secondary schools in the U.S. have acquired much of their English through ears (*aurally*) and immersion in the language.

In terms of writing performance, those international EFL students who have experienced learning English visually tend to produce pieces of writing that demonstrate grammatical understanding but limited fluency and coherence, reflecting their limited knowledge of academic forms and audience expectations, and resulting in writing that typifies Kroll's rhetoric category. This might be due to the facts that, as Leki argues, not many cultures appear to teach rhetorical patterns directly in the school settings and it seems that there are virtually no courses on writing in most countries outside the U.S. In addition, although they have learned about the structures of the language through reading rather than speaking, their practice in producing written English frequently has been limited to the classroom (1992). In contrast, the writing produced by those who learn aurally may appear to be fluent because of their fluent and comprehensible spoken language, though they may still have limited

knowledge of the structure of the language and their writing and fall into Kroll's categories of minus syntax.

Kroll's (1990b) evaluation of 100 essays written by 25 advanced ESL students at the freshman composition level at the University of Southern California (USC) seems to provide a strong empirical support for Kroll's (1990a) rhetoric/syntax split. A stratified random sampling of subjects represents the five largest foreign language groups at USC at the time (Arabic, Chinese, Japanese, Persian, and Spanish). The subjects' essays were analyzed on two dimensions: 1) the syntactic accuracy and 2) the discourse/rhetorical effectiveness of each composition. To analyze the syntactic accuracy of each composition, syntactic errors were coded to determine accuracy level of each paper. The second measure was to assign a holistic score on each essay's features of organization and coherence.

To measure the interface of syntactic and rhetorical assessments, the Spearman correlation tests were conducted using the obtained scores. The results suggest that the two scores for the compositions were not statistically correlated at the .05 level. The value of rho (the correlation coefficient) in each category is low: rho = .083 for syntactic accuracy and holistic score (rhetorical effectiveness) compared for class essays; and rho = .043 for the two scores compared for home essays. Therefore, the interpretation was that "there is no necessary relationship between syntactic accuracy and rhetorical competency in the 100 student essays" (p. 149).

2.3.3 Assessing Genre of Writing

Assessment of second language writing needs to be increasingly concerned with genre (Fulcher, 1998; Tribble, 1996; Delaruelle, 1997; Macken and Slade, 1993;

Cooper, 1998). In Fulcher's perspective, if genres are culturally conditioned and institutionalized and can be recognized and classified from their communicative purpose, assessment of L2 writing needs to account for this. He suggests that genre-based assessment is especially relevant because a learner's target performance needs can be identified with some degree of certainty. This is particularly the case in EAP/ESP settings such as academic and business English courses where the genres being taught are the ones that students need to understand and use.

In assessing L2 writing, it is common to use scales that describe the key features of the performance elicited in the writing test/task. As Pollitt and Murray (1996) point out, during the assessment process, "raters must be able to easily match the performance to the set of descriptors they use" (p. 76). Taking into account the concern for genre as stated above, Delaruelle (1997) raises the question of whether it is appropriate to use scales containing generic criteria that are applied to whatever kind of text is produced (thus possibly requiring raters to ignore features of a text that are particularly salient to a given text type) or whether there is a case for having task-specific criteria for different kinds of writing tasks.

Bhatia (1993) suggests that genre-based assessment has a number of advantages over more general approaches to language assessment. The first important one is that students bring their knowledge of the purpose, structure, and grammatical features of genres to the assessment situation/task in a way that helps them deal with unfamiliar content and vocabulary in the testing material. In addition, the relevance of the assessment that takes into account the characteristic textual and organizational features of genres to the learners' performance ensures a more accurate assessment of their ability, making the assessment a better predictor of student success.

Cooper (1998) notes that knowledge of genres of written discourse provides him with new ways of thinking about his students' struggles with writing. In his views, understanding the unique characteristics of genres allows language teachers to be able to give more productive assignments and evaluate students' writing more insightfully. He suggests that in a course of teaching writing when students are assigned to write an essay that takes a position on a controversial issue, it would be more purposeful responding to students' drafts in conference or guiding students in responding to their drafts based on knowledge of genre explicitly taught. As he puts it:

> What the student has accomplished, and might yet accomplish, can best be talked about in terms largely unique to writing that takes a position: arguing, reasoning, asserting a position, giving reasons, supporting the reasons, anticipating readers' questions and objections, and so on. Only this kind of talk can move the draft forward because so little else needs attention, even if the essay were a revision (p. 30).

Cooper also illustrates the importance of genre knowledge by contrasting two perspectives on evaluating writing, one he calls "all purpose," the other "genre-specific." These two types represent the criteria list or scoring rubric that, according to him, serves various purposes: as guides for scoring large-scale assessments of writing achievement, guidelines for the writer at work, guidelines for peer critique, as well as checklists for self-evaluation.

All-Purpose Criteria

- Focus and voice established early and maintained throughout;
- Organization effective and clearly signaled;
- Examples and details relevant to the purpose;
- Sentence structure and length varied;
- Language and tone appropriate to the purpose and readers;
- Conventions observed (p.30).

Despite the fact that many school and college writing textbooks and many instructors still prefer to rely on all-purpose criteria because it is convenient to use in judging student writing by assuming that all writing is the same, Cooper argues that they are indeed limiting and confusing because such criteria fail to inform students the possibilities of specific genres and advocates instead genre-specific criteria, which are particularly useful as guidelines for the writer, for peer critique, and for self-evaluation.

Genre-Specific Criteria: Taking a Position on an Issue

- Asserts a clear position on the issue;
- Gives specific reasons for holding the position;
- Supports each reason with personal experience, examples, statistics, or by quoting authorities;
- Provides readers with new, surprising ways to think about the issue;
- Shows an understanding of opposing views;
- Anticipating readers' objections and questions;
- Sequences the argument in a logical step-by-step way (p. 31).

As can be seen from the list, Cooper contends that these genre-specific criteria is useful in a number of ways including clearly announcing to novice writers what is to be achieved, helping teachers to focus their comments when assessing student work, and representing the criteria list for learners learning new genres through assigned reading. In a composition course, a result of genre analysis on different types of genres can be presented to students in the form of criteria lists that typify the schematic structures of each genre to help them read, understand, and approximate control of the genre more quickly. By explicitly teaching and training undergraduate students in genre-specific criteria, Chaya (forthcoming) argues that explicit knowledge of the unique characteristics of genre can develop in students as

effective metacognitive strategy used in a process of revising their drafts of argumentative essays. The result of her experimental study suggests positive learning gains as shown by a significant improvement from first to second drafts' rhetorical effectiveness of the experimental subjects' argumentative writing.

2.3.4 The Toulmin's Model

Allaei and Connor's (1991) article on using performative assessment instruments with ESL student writers suggest two important steps in establishing the validity of the assessment. They recommend that teachers first identify the text-production skills to be measured in order to present students with a clearly defined writing task that has been designed to elicit a particular type of writing, such as description, narration, or in this case, argumentation. Teachers then evaluate the criteria that constitute successful writing in the assigned genre and develop detailed scoring rubrics to evaluate the writing according to how well it has met the demands of that particular rhetorical situation. The rubrics constructed thus are task-specific, reflecting the particular content and rhetorical demands of the assigned writing task.

Various analytical scoring schemes have been suggested for assessing argumentative writing. Many of them were developed and grounded in the Toulmin's (1958) model for argument analysis. Classical rhetoric suggests that effective arguments blend appeals to logic with appeals to both the writer's credibility and the reader's emotions or affect. Toulmin's model of argument, however, focuses on logical appeals. Toulminian elements of arguments can fit in the category that classical rhetoric refers to as *logos* or logical reasoning. The Toulmin's model has six parts; three of them (claim, data, and warrant) are primary and essential, and the rest

(backing, rebuttal, and qualifier) are secondary. According to Toulmin, any argument aimed at logical reasoning processes is divided into three basic parts: the claim, the data, and the warrant. The claim is the proposition that the persuader hopes will be believed, adopted, or followed and needs support by data or evidence that gives the audience reasons for following the advice of the claim. The warrant, often unstated, is the assumption that justifies the data offered as a basis for the claim being argued. Since the literature reviewed and discussed hereinafter is limited to only three essential elements of arguments as posited by Toulmin's model, readers may consult Larson (1998) or Wood (1998) for more detail and explanation on Toulmin's model.

Yeh (1998) provides some important reasons why the model has gained prominence and is widely used to assess, teach, and study both debate and argumentative writing. The attraction of Toulmin's model is that, as he argues, the model, being built upon normal human thought processes, is specifically designed to analyze real-world argument with which people usually have had experience in the everyday argument in debate as well in composition (Fulkerson, 1996). Unlike the traditional focus of logicians that concerns with discovery of truth in established logically valid arguments rather than probabilities and thus largely ignores the type of argument common in real-world contexts, Toulmin's model suggests the possibilities for the contexts of arguments where the conclusion (claim or thesis) is only asserted to follow from the premises (support) with some degree of probability and plausibility.

In terms of pedagogical applications, the model represents the simple procedure of the basic layout of argumentation (van Eemeren et al., 1996). In assessing student writing of argument, the model as the criteria prompts the rater to consider in a step-by-step fashion whether the elements of arguments reflected in

student essays fulfill certain necessary functions. According to Larson (1998) and Yeh (1998), the rater may first locate the claim or opinion, next determine whether the writer has provided proof (data/evidence) to support the claim, and then examine the strength of the warrant (usually implied or unstated shared principles or values in the field connecting the data to the claim). If there are exceptions to the warrant, qualifiers need to be inserted. If the warrant is questionable, backing must be provided. If strong counterarguments exist, rebuttals are required. In argumentative essays presented in a problem-solution pattern produced by the subjects of this study, warrants can be detected in the introductory part that reflects the specific problematic situation of the issue, representing a basis on which the connection between the proposed solution and the (scientific) proof can be established.

In evaluating the soundness of the arguments, Toulmin, Rieke, and Janik (1979) suggest that readers begin by evaluating an argument by establishing the presence and quality of argument components. They also note that readers quickly lose faith in an argument if, for instance, the claim is ambiguous or evidence is missing. However, Toulmin cautions that an argument's quality can only be fully judged according to the standards within the appropriate field; that is, a judge may need to have field knowledge to understand and judge the quality of arguments (1958, in Yeh, 1998). While Toulmin's model is viewed to have instructional advantages in the teaching and evaluating argumentative writing in educational context, Varghese and Abraham (1998) note that students need not be specifically instructed to write to the formula of claim, data, and warrant, since the result would be stilted prose that might not achieve its persuasive goal. Rather, Varghese and Abraham's students were counseled to apply the model to evaluate their arguments. In the composition

classroom, it might be more useful to direct students during prewriting stage to consider their individual contexts of arguments (rhetorical or problematic situation) of the writing topics of their own choice before training them in genre analysis of prototypical exemplars of argumentative texts (Intaraprawat, 2002). Toulmin's elements of arguments then can be used as argument strategy taking the form of rhetorical questions that helps students develop and revise their arguments (Wong, 1988).

The validity of Toulminian elements applied in developing rating-scale descriptors is evident in a number of empirical studies (e.g., Yeh, 1998; Crammond, 1997; MacArthur and Ferretti, 1997; Ferris, 1994; Connor, 1990; Durst, Laine, Schultz, and Vilter, 1990; Thornburg, 1991; Connor and Lauer, 1988, 1985; Knudson, 1992a, 1992b; Delia, Kline, and Burleson 1979). Yeh's (1998) study attempts to validate factors that influence rating of argumentative essays in order to develop a scheme for assessing the quality of middle-school students' argumentative writing. The constructs operationalized to be the characteristics of the argument texts that the language arts teacher rate include claim clarity, strength of supporting reasons, strength of refutation to counterarguments, the extent to which the structure of the argument is developed, the degree of author's voice, and the number of conventional mistakes in punctuation and spelling. These dimensions tested are incorporated into the rating criteria and scale, describing specific aspects of an essay for the rater to examine and score. The outcomes of the study based on the results of factorial analyses assessing the effect on holistic scoring of these factors suggest that the extent of development of argument structure and the number of conventional mistakes are the two constructs that account most for the variance in holistic ratings. Both factors

explained 63 per cent of the variance in holistic ratings. Put another way, the more complete the argument structure, the higher the holistic scores, and the fewer the conventional mistakes, the higher the rating. According to Yeh, the construct 'the development of argument structure' or in Yeh's term 'Development Scale' refers to "the development, organization, focus, and clarity of the essays," (p. 139) and adherence to the 'conventions' "regards correctness of mechanics involving usage, sentence structure, punctuation, and spelling" (p. 142). The Development Scale is based on Toulmin's model of argument focusing on logical appeals and the other essential elements, which in the study are referred to as claim clarity, strength of supporting reasons, and strength of refutation to counterarguments.

Other preceding research in the field seems to provide results reinforcing Yeh's findings. One seminal work carried out by Connor (1990) suggests a scale based on Toulmin's model to be the strongest predictor of holistic scores. Connor developed a set of 11 measures based on linguistic and rhetorical features that theoretically and empirically were found to be valid and reliable indicators of writing quality of persuasive prose. The measures were used to develop the scoring criteria for teacher raters to analyze and rate argumentative/persuasive texts produced by high-school English-speaking students. In the holistic rating situation of the study, raters judged the clarity of the writer's problem statement and claim, the quality and quantity of data or reasons supporting the claim, and the quality and quantity of warrants connecting data to the claim. The results of a stepwise multiple regression analysis with an impressionistic holistic rating as the dependent variable pointed out that a set of 11 measures as independent variables account for 61 per cent of the variance in the holistic scores. More importantly, the scale based on Toulmin's model focusing on

logical reasoning was found to be the strongest entering predictor in the stepwise regression model, explaining 48 per cent of the variation in the holistic scores.

Based on the results of various studies on validating the schemes for assessing student persuasive/argumentative writing such as Connor (1987, 1990) and Connor and Lauer (1985, 1988) that have identified a number of both linguistic and rhetorical variables that are useful in predicting the overall quality of student persuasive writing, Ferris's (1994) study analyzed 60 persuasive texts written by university freshmen composition students, half of whom were native speakers and half of whom were non-native speakers of English for 33 quantitative/textual, topical structure, and rhetorical variables. The analytic scoring models based on these variables were used to examine and address the three areas of weakness in persuasive discourse of students suggested by Crowhurst (1991): inadequate content, poor organization, and stylistic inappropriateness. Using a stepwise regression analysis, two rhetorical variables, counterarguments and the added Toulmin score (claim, data, and warrant) were found to be strong factors distinguishing native speakers from non-native speakers and were also good predictors of the holistic scores given to the essays. The added Toulmin score as independent variable and as the strongest predictor of the regression model explained about 34 per cent of the variation in holistic ratings while the second best entering predictor, counterarguments, accounted for the amount of only 7 per cent of explained variance in the holistic scores.

Compared with the established validity of Toulmin-based schemes for assessing student argumentative writing as evident and strongly supported by these seminal studies discussed above, several other theoretically important variables proved to be much less significant than the Toulmin-based scale. Credibility appeals (appeals

to the writer's personal experience, knowledge of the subject/topic, and awareness of the audience's values) accounted for only three per cent of the variance in holistic scores, while affective appeals (the use of concrete, emotionally charged language and metaphors) were found to insignificant (Yeh, 1998). Regarding the influence of text coherence on holistic scorings of argumentative text quality, the results of a number of studies such as Connor (1990), Connor and Lauer (1985), and Durst et al. (1990) are mixed. Using a Bamberg's (1983) four-point scale focusing on organization, cohesion, grammar and mechanics, closure, focus, and details that describe the context for the argument, strong correlations were found between text coherence and holistic ratings (Cornor and Lauer, 1985). Durst et al. (1990) also achieved the same results as suggested by Connor and Lauer (1983) though the scale used was a modified version of Bamberg's (1983) scale.

While there appears to be a strong influence of coherence on holistic rating as indicated by these studies, Durst et al. note also that coherence is mostly likely to be correlated with logical appeals. Based on a stepwise regression analysis, their study showed logical appeals being the strongest entering predictor accounting for 53 per cent of the variance in holistic scores. When logical appeals are entered into the model, coherence accounted for only three per cent of the variance of holistic ratings. Statistically, this can be interpreted that coherence is still a strong factor but when it overlaps with logical appeals, its additional contribution to holistic rating becomes relatively small.

CHAPTER 3

METHOD

This chapter discusses the components of research method of the present study including research questions and hypotheses, subjects and subject selection procedure, research instruments, data collection and analysis, and statistical treatment.

3.1 Research Questions and Hypotheses

This instructional study sets out to investigate the effects of using a process-genre approach in the teaching of argumentative writing on the students' writing performance. The research questions are:

1. In what aspect(s) of the subjects' written argumentation can the use of genre-based approach emphasizing explicit teaching of rhetorical elements of argument prove to be effective in facilitating the student writers to produce an effective argumentative essay?

2. In what way(s) do the experimental group's students find the instruction useful in helping them to write an effective argumentative essay?

The related hypotheses to be tested are:

1. There will be a significant difference between the experimental and control groups in their first- and final- argumentative essay drafts' gain scores on an overall impression of their essays, and on the three separate rhetorical qualities

of written argumentation including claims, reasons, and rebuttals to counter-arguments.

2. The experimental group's students will find the instruction helpful as reflected by a positive attitude and feedback toward the instruction.

3.2 Subjects

The participants were 55 non-randomized L1-Thai undergraduates enrolled in course 203305 titled English V: Argumentative Academic Writing in Trimester 2 of the academic year 2003 at Suranaree University of Technology (SUT), a state university located in the northeastern part of Thailand. The subjects consisted of students from varied specialized fields of engineering and health studies at the university.

These non-randomized subjects were chosen and included in the study based on the different instructional treatments they had been exposed to over the 12-week course duration. The two classes of approximately 30 students each, taught by two teachers using two different approaches to teaching argumentative writing were selected for comparing the effects of each approach on the qualities of written arguments produced by the subjects at the end of the course. The experimental group students were taught by a teacher who used a genre-based method or direct instruction of rhetorical features of argument genre, and the other controls (comparison group) were exposed to the method focusing on the writing process alone.

The subjects' characteristic and EFL writing proficiency can be inferred from the subjects' responses to some of part 1 items of a questionnaire (see Appendix C) and a writing pretest, respectively. The writing pretest, however, was administered to

roughly evaluate the subjects' writing ability only, that is, to indicate how good or poor they were in EFL writing before the instruction began. The administration of the writing pretest and questionnaire was intended to get round one weakness of a static group comparison design of no examination of pre-experimental equivalence of groups by providing as much an ethnographic detail as possible of the two groups of subjects (see 3.3.1 for further discussion).

As pointed out by the questionnaire responses, almost all of the subjects admit that EFL writing has been quite difficult and very problematic for them particularly in the areas of grammar and vocabulary, and all of them have never been taught to write argumentative essays before they attend English 5. The typical types of writing assignments they have experienced so far are mostly paragraph writing responding to assigned topics or writing summaries of assigned readings. The average length of the writing ranges from 1 to 3 paragraphs but not more than two pages at most.

The resulting written products of the writing pretest administered before the instruction began were likely to reinforce such responses. The subjects' written responses to the argumentative writing prompt showed that many students seem to lack the ideas of how a format and organization of what standard English essay should look like. It is, however, surprising to see that the respondents did not state such a problem to be an important one in comparison with the language problem. The indication may contribute to the possibilities that they were overwhelmed by the problems of lack of necessary grammatical knowledge and of inadequate vocabulary size to be used in expressing their intended ideas through writing, and/or being unaware of the difference of rhetorical conventions between their L1 Thai and L2 English written discourse. A series of studies by Ward (2002, 2001, & 1999) on the

vocabulary size of SUT engineering undergraduates that reveals the association between SUT engineering undergraduates' reading difficulty and students' tendency to avoid reading English academic texts because of their insufficient vocabulary size can, to a large extent, support this assertion.

Thus, it would be safe to say that at the point where the instruction was about to begin, both subject groups' writing proficiency and problems were not varied. Typically, of these two groups of students, most seem to have limited knowledge of rhetorical conventions of English argumentative written discourse as well as argumentation skills while some also had serious difficulties concerning syntactic and vocabulary knowledge. Based on these elicited data, it was apparent that when assigned to write an argumentation, the subjects would construct their own patterns in writing arguments.

3.3 Design

The present study was quasi-experimental (a static group comparison) in its nature. The components of the design can be described in terms of the subject and data, the treatment, and observation and measurement of the treatment (Selinger & Shohamy, 1989). Using conventions established by Campbell and Stanley (1963), the quasi-experimental design with control group of this study can be symbolized as follows:

Group 1 (Experimental group)-----X--------O1

Group 2 (Control group)----------------------O1

As represented by the symbols, this research concerned with studying the effects of one particular method of teaching argumentative writing (represented by X)

to EFL adult learners in institutional settings (experimental group) on their argumentative writing ability (group 1's O1), and comparing it to those of control group representing the same population as the experimental subjects (group 2's O1) but receiving an instruction different from that of the experimental group. Figure 3.1 provides a flowchart of this study's design.

Figure 3.1 The Study's Quasi-Experimental Design

A sample of 55 subjects participating in the study

↓

Two intact classes of subjects

Experimental group:
28 subjects

Comparison group:
27 subjects

↓

The independent variables are manipulated.
(Two different instructional approaches)

↓

The subjects perform the argumentative writing task.
(Argumentative essays = the dependent variable)

↓

The dependent variable is evaluated:
Raters read and score the essays against the constructed rating scales.

↓

Measures of central
tendency and variability
calculated
(means and standard deviations)

Measures of central
tendency and variability
calculated
(means and standard deviations)

Between-subjects
comparison (Statistical Tests)

3.3.1 The Type and Number of Groups

The study utilized two intact (non-randomized) groups of students already existed at the time of undertaking this research since it was not possible to randomly select and reassign them to form groups in formal institutional settings. Although the subjects were not randomly selected, they could be considered randomized ones on the basis that each of them was free to choose a class he or she wanted to be in from a number of available classes without knowing in advance who was going to be a class teacher when the class was about to begin.

Statistically, sizes of the subjects of each group (n = 30) are considered a large sample size for test statistics. This means that the subjects were likely to represent a normal distribution of the population or a range of writing abilities that could be expected from the whole student population registered in the same course in the same trimester. Each of the two intact groups consisted of about 30 students more or less as reported by the university registration list.

However, based on the reason that the data (the subjects' essays) to be collected and analyzed were viewed to be the outcome of the hypothesized effects of the two different teaching methods and specially constructed studying material, some subjects' written products were not valid to be included in the study if the individuals failed to meet one or a combination of the following inclusion criteria including 1) being absent from class during the critical instruction period (period of instruction hypothesized to have significant impact on the subjects' written products); 2) not submitting writing assignments on time or not having a complete set of written drafts and; 3) an essay that was rated 0 on a holistic rating scale indicating its inadequacy for

a further analysis (analytic scoring), and not included in hypothesis 2 testing. These criteria were applied to both groups.

3.3.2 The Treatment

Two intact classes participated in this study. Both groups of subjects were exposed to two different instructional approaches and teaching and learning activities for 36 hours contact (3 hours per week for 12 weeks). The instructional treatments utilized in this study were similar for both groups of subjects in their emphasis on prewriting, drafting, revising, and editing activities within the writing process. To achieve control, what differed between the treatments received by both groups of subjects, however, was the types of teaching and learning activities in which the students participated, the focus of instruction on the context of writing or on the structure of text in the writing process.

For the experimental class, the instructional treatment using an integrated process-genre approach was operationalized into the following sequence of stages: (1) modeling; (2) joint negotiation of text; and (3) independent construction of text. These stages are shown in the circle diagram (Figure 3.2).

The teaching and learning cycle consists of different phases of classroom interaction and learning support which the teacher and learners go through so that learners gradually gain independent control of the genre. The cycle was flexible that each stage can be repeated as needed if students need revision or further practice in order to progress. The between-groups difference in learning gains as a result of the two different types of instructional treatments can be explained by the presence of these three stages of the teaching and learning cycle experienced by the experimental group, ceteris paribus.

Figure 3.2 The Teaching and Learning Cycle (based on Feez and Joyce, 1998)
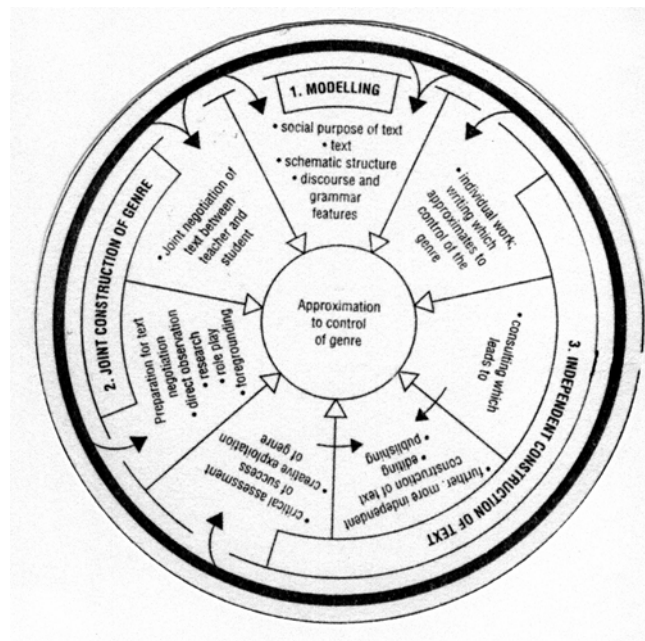
The first modeling phase involves context building and modeling of the text. The first important foundation step in this phase, context building, was to raise the students' consciousness of the contextual features of the genre in terms of the social context (rhetorical situation including purpose, audience, and occasion) in which the genre is used and the purpose it serves. This was followed by modeling, in which the students were exposed to texts generically representing the genre and a discussion on how the form is functional, how the information is organized (schematic structure), and the distinctive language aspects (lexico-grammatical features) which realizes the text. Phase two involved joint construction of the text including teacher-student discussion of the genre where the teacher guided the class composition through questioning, discussing, editing whole class construction, and scribing onto the overhead. In the third and final phase, students individually constructed the genre, basing their drafts on research work on the context and the topic of their writing, and consulting with peers and the teacher as needed.

The control class, on the other hand, learned and practiced writing within an environment that emphasized the writing process with a particular focus on audience and purpose of writing. Though the control subjects were not explicitly introduced to text structures as given to the experimental group, they were sensitized to the situational context of the argumentative text through the use of think-sheets in terms of prompts and questions focusing on audience and purpose of writing during the writing process. These prompts were used as instructional as well as modeling strategies in introducing to the control group the different strategies and activities appropriate to various aspects of writing during prewriting, drafting, revising, and editing.

Through the span of the taught course, the subjects were encouraged to produce and revise several drafts of their argumentative writing according to the process writing instruction given to both instructional groups. It should be noted, however, that the experimental group was required to submit their first drafts at the end of week 6 of the 12-week course. The control group was also expected to finish and submit their drafts at the same time but the submission deadline was not as strictly applied as in the experimental group.

According to the hypothesized effects of using a genre-based method of explicitly teaching students to analyze rhetorical pattern and organization of the argumentative genre as early as in the prewriting stage, both instructional groups' first drafts were collected for scoring and treated statistically to investigate the research question 1. While not overlooking the pedagogic potential of the process writing instruction received by both groups of subjects that could give all the subjects a fair

chance of success in their writing performance, both groups' final drafts were also collected for scoring and the gain scores treated and compared statistically.

3.3.3 Observation and Measurement

3.3.3.1 Observation

For the sake of thorough understanding of the classroom teaching practice adopted by each group's teacher, the researcher as observer took a role of an insider by sitting in all class meetings of both groups and did note-taking to record a detail account of classroom administration and teaching practice, in-class activities and writing assignments, as well as the lesson plans, for instance. The data obtained from the observation were used mainly in supporting the discussion of the results of data analysis and also in drafting up questionnaire items.

3.3.3.2 Measurement

To measure and compare the effects of the two different instructional approaches, all the subjects' argumentative essays were collected and rated by three raters against the specially constructed rating scales. The researcher was a team leader and one rater. The other two were experienced EFL teaching professionals; one teaches 7 to 12 graders at a local secondary school, and the other teaches English to undergraduates at SUT. The ratings consisted of two sessions with the holistic scoring done first, followed by the analytic session. Three raters were used for the holistic reading, while only two of them were available for the analytic scoring since the other one was preoccupied with busy teaching schedule and work load. Since the raters know the teachers of both classes, the collected essays were anonymously mixed to

prevent the possibilities of the halo effect on the part of raters, which would seriously reduce the validity of scores assigned to essays (Hopkins, 1998). All assigned scores were treated statistically to see if there were any significant difference between the two groups of subjects in terms of their average holistic and three separate analytic scores.

To obtain feedback from the subjects as to how they perceived of the teaching and learning, a questionnaire as a supplementary tool was administered in the last class meeting of the course. The responses received were used to triangulate the result of data analysis or to support the research findings.

3.4 Research Tools

The two types of instruments specially constructed and used were the two kinds of rating scales (holistic and analytic), and a retrospective questionnaire. The rating scales were used as a major tool to find the answer to the first research question and its related hypotheses. The questionnaire was used for eliciting additional retrospective information from the experimental subjects concerning their EFL writing experience, perceived difficulties and advantages in learning and writing argumentative essays, and comments on the usefulness of the instruction. The responses from experimental subjects were used to determine whether they found the instruction facilitative in helping them to write argumentative essays.

3.4.1 The Rating Scales

The two types of ad hoc scoring criteria, holistic and analytic rating scales were constructed and used by the raters for evaluating the subjects' scripts. The rating

scales define explicitly the criteria against which raters read and score the subjects'

essays. Their development and characteristic features were discussed below in order.

3.4.1.1 Scale Development

The scale and its descriptors were empirically generated following the

data-based approach to scale construction advocated by North and Schneider (1998),

as opposed to the mastery approach suggested by Bachman and Palmer (1996) (see

2.3.1). To ensure the validity of the scale, the following procedure was undertaken in

incorporating into the scale descriptions the abilities that the subjects were expected to

perform as they were taught in writing their argumentative essays. The scale

descriptions were done empirically through gathering sample scripts from students at

all relevant levels of proficiency, examining the scripts for textual features of

performances at different levels, and seeking expert judgements from a number of

English 5 instructors familiar with the performance levels of students for defining the

characteristics that differentiate the samples. The sample scripts collected for the

empirical development of the rating scale were those essays submitted for scoring at

the end of trimester 1/2003.

3.4.1.2 Holistic Rating Scale

A holistic rating scale (Appendix A) outlines the criteria for scoring the

subjects' scripts on a basis of the overall impression of an essay. The scale contains

descriptors of syntactic (language) and rhetorical qualities of six levels of writing

ability with the lowest scale 0 to be awarded on the script that is inadequate for

assessment or fails to meet the least requirements expected by the course. The six

levels of the scale were adopted a priori, by deciding in advance considering the most

common numbers of levels used in well-researched large-scale assessment programs such as TOEFL/TWE or Michigan Writing Assessment Scoring Guide, and by consulting with the researcher's supervisor for the range of performance that can be expected of the population of English 5 student writers.

The use of rhetorical and language descriptions for each level was based on the framework for assessing ESL writing of non-native speakers proposed by Kroll (1990a) which demonstrates that ESL/EFL writing proficiency can be broadly divided into "plus-syntax and minus-syntax" and "plus-rhetoric and minus-rhetoric." As the term 'rhetoric' (usually refers to content, organization, and development) of argumentative genre is unique in its own right and thus deserves a careful attention in defining the 'rhetoric descriptor,' simplified Toulminian elements for argument analysis were applied for writing up the rhetorical qualities of each level of the scale. The descriptions appeared in the scale implicitly refer to such notions as claims, reasons, and refutations to counter-arguments (in 'rhetorical control' category), and the facility to use the grammatical system of standard edited English in such categories as sentence structure, word form, word order, verb tense (in 'language control' category).

Although there would be cases that students' writing performance may not fulfil both 'rhetorical' and 'language' criteria of the same level of Kroll's framework, holistic scoring was still useful and adopted mainly as criteria for judging the writing as a norm-referenced procedure for scoring, sorting and ranking pieces of writing of each group of subjects based on a reader's reaction to the impression of the text as a whole, not to the parts of it as in analytic scoring.

3.4.1.3 Analytic Rating Scale

An analytic scoring rubric was also used in the present study to cope with the downside of holistic scoring in not providing useful diagnostic information about a person's writing ability. As illustrated in Appendix B, descriptions of each level of rhetorical descriptors of the holistic scale were rearranged into three quality aspects of argumentative writing consisting of the three separate Toulmin elements of argument (claims, reasons, and refutations to counter-arguments). Since the study concerns the possible hypothesized effects of the instruction on the qualities of claims, reasons, and rebuttals to counter-arguments of the subjects' essays as a direct influence of the independent variable (a direct instruction of rhetorical patterns), the other aspect, language use, was omitted and the raters used only six levels of rhetorical descriptions for judging the scripts. To score analytically, the raters had to read through the errors of syntax and attend only to what Kroll termed 'pure discourse features' of an essay or the level of organization, coherence, and discourse fluency of argumentative written discourse features as described above. In other words, the essays were being scored as if they had no grammatical errors.

The conduct of analytic scoring can be justified on the notions that holistic scores are not always easy to interpret, as raters do not necessarily use the same criteria to arrive at the same scores, and as a single score does not provide useful diagnostic information about the subjects' writing ability (Weigle, 2002, Reid; 1993; and Hughes, 1989). For example, a certain script might be given a 4 on a holistic scale by one rater because of its rhetorical features (content, organization, and development), while another rater might give the same script a 4 because of its linguistic features. In addition, a single score 4 does not allow raters to distinguish

between various aspects of writing. In this study, such aspects particularly refer to rhetorical elements of argumentation including development and organization of claims or theses, reasons supporting the claim, and rebuttals to counter-arguments that need to be assessed, letting alone the language use aspect.

Therefore, the use of analytic scale allows the researcher to study and compare the effects of the two different instructional methods on the qualities of claim, reasons, and rebuttals to counter-arguments of each script of the two groups' students which could not be made possible if using the holistic scale alone.

3.4.1.4 Rater Training

To maintain high consistencies in assigned scores by different raters (inter-rater reliability) from which appropriate inferences about a range of writing ability of the subjects can be made with satisfactorily level of accuracy, three raters were trained in operational scoring of the selected sample scripts using the constructed rating scales and benchmark scripts that exemplify the different points on the scale. In the present study, three raters were used and trained before scoring the subjects' argumentative essays, though empirical studies have consistently shown acceptably high scorer reliability (inter-rater reliablity) when writing (each student's essay) is scored four times or by four raters (Hughes, 1989).

In the first stage, the first set of benchmark scripts with the appropriate scores indicated was given to raters to familiarize them with the scale and to instantiate certain features of the rubrics. The trainer (researcher) used these scripts to describe for the raters what is meant by phrases in the rubric. Questions raised by the raters were discussed until agreement achieved. Once the raters felt comfortable with

the scale as instantiated in the first set of scripts, in the second stage, a set of ten scripts of different scales without scores indicated was given for the raters to practice reading and scoring against the scale. Once the practice session finished, all raters discussed on the assigned scores to reach the finalized scores. Additional training was also given to the raters to bring their scores into alignment with each other in the group. The training process was the same for both holistic and analytic scoring practices.

A Statistical Package for the Social Sciences program (SPSS Version 11.0) was used for reliability analyses of the three raters' assigned scores from both practice-run scoring sessions and a cross-tabulation of pairs of raters. The inter-rater reliability coefficient alpha of the holistic and analytic scoring training sessions were 0.87 and 0.92, respectively, which are adequately high when compared with a normally accepted level of 0.80 or greater. Following White's (1984) criteria, an average success of the whole scoring-practice session was achieved as the ratings of more than one point apart by any pair of raters were not more than 10% (cross-tabs), which indicated that they were ready for the real scoring session, and that, in turn, reflected the validity of the constructed rating scales.

3.4.2 Questionnaire

A retrospective questionnaire was administered anonymously before the end of the course to elicit additional feedback from the experimental subjects with regard to their English writing experience so far as well as in SUT, perceived advantages and difficulties about the way they were taught, and comments they might have about the course.

The questionnaire was originally generated in English but translated into Thai at the time of administration for the sake of full comprehension of the respondents. Based on the limitation that the questionnaire items (specifically those in Part II) were designed to incorporate elements hypothetically established to be the merits of the genre-based approach to teaching argumentative writing, the questionnaire was administered to the experimental group only.

The questionnaire items were grouped into three parts according to the types of responses to be elicited by the questions. A set of questions in Part I was designed to seek to know from the respondents their EFL writing experience before attending English 5 and their attitudes toward English language learning particularly writing. Part II contained questions designed to match the hypothesized effects of the instruction given to the experimental group. The final part of the questionnaire, Part III, contained one open-ended question to elicit the experimental subjects' opinions, suggestions, or even complaints they would like to express concerning the conduct of the class, the way they were taught, or the writing problems they encountered during the course, for example.

In developing the questionnaire, reliability analysis of its question items was not conducted or Cronbach's alpha coefficient which determines a good estimate of reliability based on internal consistency not calculated since it was not within the scope of this research. Instead, appropriate actions were undertaken to assure validity of items and reliability of responses to the items. This is especially the case of items in Part II. The steps involved resembled to those conducted in constructing the scoring rubrics in that it was done empirically. To ascertain the validity and reliability of the items in the sense that it could elicit the intended information that represent a

true range of the subjects' characteristics, the items were built upon the elicited range of categorized responses from eight experimental subject volunteers to a set of open-ended questions based on classroom observation data obtained by the researcher. It is important to note that in developing a set of items in Part II which was class-specific, the researcher had sought expert judgement by working closely with a highly experienced writing teacher of the experimental class whose critique was taken into account in drafting up the questions and making necessary modifications. Once the validity of the questionnaire was achieved, it was of equal importance to assure that the elicited responses were, at least to a certain extent, reliable or generalizable. By piloting the questionnaire to the same group of volunteers, the researcher collated and analyzed their preliminary responses. The volunteers were also interviewed if they could fully understand the questions, and indicate the unclear items and response choices. Such interviewed data were used in making appropriate modifications to those ambiguous items in making a final draft of the questionnaire.

3.5 Data Collection and Analysis

The data used for the analysis were 55 essays produced by the subjects at the end of the course. Copies of all essays were distributed to all three raters for reading and scoring. All papers were mixed and treated anonymously so that the raters did not know to which groups each individual essay belonged. The procedures of data analysis consisted of four major steps: operational scoring sessions, reliability analysis, testing of hypotheses, and analysis of questionnaire responses.

3.5.1 Operational Scoring Sessions (Holistic and Analytic ratings)

To analyze the collected scripts, the holistic scoring was done first, followed by the analytic scoring. In the holistic scoring session, each rater independently read and assigned a single score to a script based on the overall impression of the script. Each script was read quickly and then judged against the holistic rating scale. In the analytic scoring session, each script was rated on the three separate aspects of argumentative writing features outlined in the analytic rating scale by assigning three separate scores for each of the three features. The scores given were noted down on separate score sheets, which were to be collated by the researcher to look for discrepancies. Common practice on a six-point scale is that the ratings of more than one score point apart (e.g., one score of 3 and one of 5) are considered discrepant and must be resolved. In the cases where discrepancies of assigned scores awarded on the same piece of writing were more than one, appropriate measures were adopted to resolve the cases. In the present study, discrepancies in rater scores were to be resolved by a series of discussion between the raters on a case-by-case basis to settle the disagreement and provide the finalized scores. Assigned scores on all scripts were collected and treated statistically to test the established hypotheses. All resolved scores were summarized and reported as descriptive statistics.

3.5.2 Calculating Inter-Rater Reliabilities

Because of the possible subjectivity associated with rater scores, the assigned scores needed to be validated for reliable statistical inferences to be made when testing the hypotheses and for determining the degree of success of the overall

ratings. To do so, consistency in assigned scores (scorer reliability) within and between raters had to be examined.

To assure consistency within individual (intra-rater reliability or the tendency of a rater to give the same score to the same script on different occasions), each rater was recommended to do multiple scoring.[5]   That is, each rater scored each essay two times on two different occasions and decided the final score. The second step was to ensure there was consistencies in scores given by the different raters (inter-rater reliability or the tendency of different raters to give the same scores to the same scripts).   To do so, all the final scores assigned by the three raters were collated, resolved for discrepancies, and calculated using Pearson Product Moment Correlation (r) formula. The higher the value of the correlation coefficient (r) approaching 1, the stronger the agreement between the raters or their assigned scores, indicating the more reliable the statistical inferences can be made from the scores.   Alternatively, the overall success of the rating sessions can also be determined by cross tabulating the rater scores. By adopting White's (1984) suggestion, an average reading using a six-point scale will have 7-10% of the ratings more than one point apart, while in an excellent reading, only 5% of the ratings will be discrepant.

To determine the degree of success of the overall ratings, a SPSS Version 11.0 was utilized for conducting reliability analyses (inter-rater reliability) and cross-tabulation. A threshold level of reliability coefficient alpha was expected to be 0.80 or greater. For cross-tabulation analysis, the possibilities of discrepancies were expected not to exceed 10%. The results of statistical procedures used to compute inter-rater reliabilities were summarized and interpreted according to the statistical methods used.

3.5.3 Hypothesis Testing

Upon completion of the rating sessions, all assigned scores were statistically tested for making inferences about the writing ability of both groups of students as a result of the two different instructional treatments. T-test statistic was utilized to test a set of hypotheses of research question 1. A two-tailed t-test was used to test whether the average holistic gain scores of both groups were significantly different or not. Also, a one-tail t-test was used to determine if each of the three separate mean scores gained by the experimental subjects were significantly higher than those of the control group.

3.5.4 Analysis of questionnaire responses

The main aim of administering the questionnaire was to obtain feedback from the students (experiment group) about the usefulness of the instruction, its advantages and disadvantages, and their comments on the conduct of the course. The responses given by the respondents were counted for frequency per question items and converted into percentage. The written responses from Part III were classified and summarized. All analyzed responses were used for reporting the students' feedback and for providing support to the research findings.

CHAPTER 4

RESULTS

To investigate the research questions, the established hypotheses were tested statistically. The results of the study were presented and discussed in four sections starting from the descriptive statistics, reliability measurement of the data, testing of hypotheses, and questionnaire results.

4.1 Descriptive Statistics

The descriptive statistics of the ratings are shown in Tables 4.1 to 4.4 – the first two for the holistic ratings (of final drafts only), and the other following two for the analytic ratings (of both first- and final drafts).

As Table 4.1 displays, the means of gain scores and standard deviations (SD) of both experimental and control groups were very close to each other with merely a small difference of .28 in the holistic mean scores, and of .142 in the values of SD. The mean score on the experimental-group writing performance, however, was higher than that of the control group.

Table 4.1 Holistic Ratings – Means and Standard Deviations

	Holistic Ratings	
	Experimental Group	Control Group
N	28	27
Mean	*2.28*	*2.00*
Std.Deviation	1.451	1.593

Table 4.2 illustrates the distribution of holistic scores gained by each group of subjects expressed in frequency and percentage. By comparison, it can be seen that in percentage the experimental group gained about .5 to 1 scale score apart from the control group's gain level, concentrating at the scales of 2 to 3. For the control group, the scores seem to distribute evenly from the scale 1 to scale 2.5, with a relatively more concentration at scale 0.5. For the scales higher than 2.5, the control group's performance was not as good as that achieved by the experimental group with only 1 control student scored 3, compared with 4 experimental students receiving the same score. At the levels beyond scale 3, only one student from the experimental group scored 3.5, and the other from the same group scored 4.

Table 4.2 Holistic Ratings – Distribution of Scores by Group

Holistic Ratings		Experimental Group N = 28		Control Group N = 27	
		Freq.	(%)	Freq.	(%)
Resolved	Score 0	--	--	--	--
Scores	Score 0.5	1	*(3)*	7	*(26)*
	Score 1	1	*(3)*	4	*(15)*
	Score 1.5	3	*(11)*	5	*(18)*
	Score 2	8	*(28)*	5	*(18)*
	Score 2.5	9	*(32)*	5	*(18)*
	Score 3	4	*(14)*	1	*(3)*
	Score 3.5	1	*(3)*	--	--
	Score 4	1	*(3)*	--	--
	Score 4.5	--	--	--	--
	Score 5	--	--	--	--

Table 4.3 shows the mean scores and standard deviations of analytic ratings of first- and final drafts. For between-groups comparisons of scores on the three qualities of written argumentation (Claim, Reason, and Rebuttal), the experimental group

gained higher mean scores than those of the controls across the board for both first- and final drafts, though the between-groups differences of mean scores of both types of scripts were not much different from each other, clustering around .5 level or only half a scale score. For within-groups comparisons, the amounts of difference in the means of control group across the three measurements of written argumentation were relatively slightly higher than that of the experimental group, especially the mean scores on the qualities of reasons and rebuttals to counter-arguments.

Table 4.3 Analytic Ratings – Means and Standard Deviations

		Analytic Ratings			
		Experimental Group N = 28	Control Group N = 27	Experimental Group N = 28	Control Group N = 27
		1st Draft		Final Draft	
Claim	Mean	2.25	1.74	2.43	1.93
	Std. Deviation	.928	.656	.997	.781
Reason	Mean	2.32	1.89	2.64	2.44
	Std. Deviation	.905	.892	.989	.974
Rebuttal	Mean	2.04	1.44	2.18	1.63
	Std. Deviation	.881	.847	.983	1.006

Table 4.4 displays the distribution of three separate scale scores of analytic ratings by group by type of data (first- and final drafts). For between-groups comparisons, the experimental students seem to outperform the control students in all areas of measurement as indicated by the distribution of gain scores of both groups. Also, the concentration of scores of the control group was in the range lower than that of the experimental group, particularly the distribution of gain scores on the qualities of written claims, and rebuttals to counter-arguments.

Table 4.4 Analytic Ratings – Distribution of Scores by Group

Analytic Ratings (resolved scores)		Experimental Group N = 28		Control Group N = 27	
		1st Draft	Final Draft	1st Draft	Final Draft
		Freq. (%)	Freq. (%)	Freq. (%)	Freq. (%)
CLAIM	Score 0	--	--	--	--
	Score 1	6 (21)	5 (18)	10 (37)	9 (33)
	Score 2	12 (43)	11 (39)	14 (52)	11 (41)
	Score 3	7 (25)	7 (25)	3 (11)	7 (26)
	Score 4	3 (11)	5 (18)	--	--
	Score 5	--	--	--	--
REASON	Score 0	--	--	--	--
	Score 1	5 (18)	3 (10)	11 (41)	5 (18)
	Score 2	12 (43)	10 (35)	9 (33)	9 (33)
	Score 3	8 (28)	10 (35)	6 (22)	9 (33)
	Score 4	3 (11)	4 (14)	1 (4)	4 (15)
	Score 5	--	1 (3)	--	--
REBUTTAL	Score 0	--	--	4 (15)	3 (11)
	Score 1	8 (28)	8 (28)	9 (33)	10(37)
	Score 2	13 (46)	10 (35)	12 (44)	9 (33)
	Score 3	5 (18)	7 (25)	2 (7)	4 (15)
	Score 4	2 (7)	3 (10)	--	1 (3)
	Score 5	--	--	--	--

It is quite clear from the reported descriptive statistics that the experimental group was likely to perform better than the comparison group as suggested by the resolved scale scores from the holistic and analytic ratings. Nonetheless, to examine whether the experimental group actually did so, the scores of both groups need to be treated statistically. The next section reports the results of reliability analyses of the data (rater scores), followed by the results of statistical treatments of the data. The results of both were discussed with interpretation.

4.2 Scorer Reliability

Rater scores derived from the holistic and analytic scoring sessions were used as the numerical data for statistical treatments to determine whether the experimental group outperforms its counterpart, the control group, or not. However, before conducting statistical tests, it was necessary that the data (scale scores) had to be checked for consistencies between raters in their assigned scores. This is to make sure that the raters were consistent and the ratings met the consistency criteria expected. Three statistical formulas were applied to calculate the degree of rater consistencies (inter-rater reliability coefficients) or to measure the extent to which the raters 'go together' in their assigned scores. The formulas used include Pearson Product-Moment Correlation, One-Way Analysis of Variance, and Cross-Tabulation. The results of which will complement one another for deriving the conclusion whether the scores were reliable and appropriate for statistical inferencing or not.

4.2.1 Correlation Analysis

Pearson correlation ($\underline{r}$) was applied to measure the degree of relationship between two sets of rater scores. The greater the $\underline{r}$ value closer to 1, the higher the likelihood that the two sets of scores are similar or close to each other. To measure the strength of rater agreement, r squared ($\underline{r}^2$) or coefficients of determination were also computed to examine the extent to which the variance (the variability of scores around the mean) in one set of rater scores can be accounted for by the other.

For holistic ratings using three raters, correlation coefficients and coefficients of determination were computed for all three possible pairs of raters, and three pairwise comparisons between each rater scores and resolved scores were also

conducted. As Table 4.5 shows, correlational analyses were significant for all pairwise comparisons (column labeled 'r'). To determine how much of a set of scores was overlapped by the other for any pair of correlational analysis, the $\underline{r}$ value for each rater pair was squared. In comparison, the values of $\underline{r}^2$ among the three rater pairs were much lower than those of each rater paired with the resolved scores except that of rater 2. Using the Spearman Brown Prophecy formula, the inter-rater reliability coefficient of the three sets of rater scores was .74, suggesting an adequately strong correlation among all rater scores.

Table 4.5
Holistic Ratings – Correlation Coefficients (r) & Coefficients of Determination (r^2)

(N = 55 essays)	r	r^2
Rater1—Rater2	.66	.44
Rater1—Rater3	.79	.63
Rater2—Rater3	.57	.33
R1-Resolv.	.92	.85
R2-Resolv.	.73	.53
R3-Resolv.	.84	.71

Note: Inter-rater reliability coefficient of 3 raters equals to .74

For analytic ratings using two raters (Table 4.6), the computed inter-rater reliability coefficient of .68 with its associated $\underline{r}^2$ of 46 % could be considered moderately high, indicating the ratings overlap by as much as 46 per cent. For correlational analyses of separate analytic scores, a greater increase in the agreement between each rater scores paired with the resolved scores results in the higher overlap of such pairs of scores. For a correlation between 'Rater 1 and Rater 2,' the values of $\underline{r}$ and $\underline{r}^2$ were not quite different from each other.

Table 4.6
Analytic Ratings – Correlation Coefficients (r) & Coefficients of Determination (r^2)

All Scores Combined	N	r	r^2
Rater1—Rater2	165	.68	.46

Separate scores (N = 55 each)	N	Claim		Reason		Rebuttal	
		r	r^2	r	r^2	r	r^2
Rater 1 – Rater 2	55	.65	.42	.66	.44	.68	.46
R1—Resolv.scores	55	.75	.56	.83	.69	.75	.56
R2—Resolv.scores	55	.92	.85	.86	.74	.84	.71

4.2.2 Tests of Homogeneity of Variances

For both holistic and analytic rating sessions, Levene test and ANOVA were applied to measure whether the mean rater scores were significantly different or not. Since the follow-up tests involving Post Hoc Tests and Homogeneous Subset can be run on more than two variables (more than two sets of rater scores), they were conducted for the holistic scoring only to determine which pair(s) of raters were significantly different from each other in their assigned scores.

As shown in Table 4.7, for holistic ratings, the p values of Levene test and ANOVA (.95 and .62, respectively) indicate no significant differences among the three sets of rater scores. In contrast, for analytic ratings, the computed ANOVA (p = .022) suggests a significant difference between the two raters' scores. In the sub-table that follows, an ANOVA for analytic ratings indicates no significant differences between the rater scores on the qualities of reasons, and rebuttals to counter-arguments (p value = .547 and .131, respectively) except the scores given on the quality of claims (p = .046).

Table 4.7 Levene Test and ANOVA

Rater Means	Holistic scoring N = 55	Analytic scoring N = 165
	Sig. (alpha = .05)	*Sig. (alpha = .05)*
Levene statistic	*.95*	*.069*
ANOVA	*.62*	*.022*

Analytic scoring (Rater 1—Rater 2)	*N*	*Sig. (.05)*
Claim	55	*.046*
Reason	55	*.547*
Rebuttal	55	*.131*

To conduct multiple comparisons, Tukey HSD with equal variances assumed (as suggested by the Levene test statistics of .95 and .069 shown in Table 4.7) were used to compute the p values of the post hoc tests and homogeneous subsets. The computed p of the post hoc tests (Table 4.8) suggest no significant difference between the three pairs of raters' sets of scores with the probability values for each pair less than .05 with 95% confidence interval for each pair containing zero values.

Table 4.8
Post Hoc Tests (Multiple Comparisons using Tukey HSD) (N = 55)

Pair of Rater	*Sig.*	95% Confidence Interval	
		Lower Bound	Upper Bound
Holistic scoring			
Rater 1 – Rater 2	*.84*	-.57	.93
Rater 1 – Rater 3	*.60*	-.44	1.06
Rater 2 – Rater 3	*.92*	-.63	.88

Table 4.9 displays sets of mean scores that were not significantly different from each other. In the first column, the rater scores were listed in order from that with the smallest mean (scores assigned by Rater 3) to that with the largest mean

(those assigned by Rater 1) with that of Rater 2 in between. The next two columns to the right list the actual means of each rater scores, grouped into two subsets. With the empty subset 2 (column labeled 2), all rater mean scores were grouped in subset 1 (column labeled 1), indicating that such three mean scores formed a homogeneous subset and were not significantly different from each other.

Table 4.9
Homogeneous Subsets (Tukey HSD) (N = 55)

Rater scores	Subset for alpha = .05	
	1	2
Holistic scoring		
3	*3.95*	--
2	*4.07*	--
1	*4.25*	--
Sig.	*.60*	--

4.2.3 Cross-Tabulations

By cross-tabulating rater scores, a number of discrepancies were tallied and converted into percent to indicate how often the raters agree and to check whether the established allowance criteria were met or not. As shown in Table 4.10, both scoring sessions using a six-point scale had about 6 to 7 % of the ratings more than one point apart and, according to White (1984), were considered an average reading. For holistic ratings, frequency of discrepancies ranging from about 5 to 9 % for each cross-tabulation of rater scores suggests an average reading with an almost excellent reading for the pair of Raters 1 and 3.

Table 4.10
Cross-tabulations (discrepancy allowance = 10 %)

Pair of Rater	Discrepant Scores		
	Frequency	N	%
Holistic scores (3 raters)			
Rater 1 & Rater 2	4	55	*7.27*
Rater 1 & Rater 3	3	55	*5.45*
Rater 2 & Rater 3	5	55	*9.09*
Total	12	165	*7.27*
Analytic scores (2 raters)			
Claim	6	55	*11*
Reason	2	55	*3.63*
Rebuttal	3	55	*5.45*
Total	11	165	*6.6*

For analytic ratings using two raters, cross-tabs of the scores on the qualities of reasons, and rebuttals to counter-arguments indicate an excellent reading with only 3 and 5% of the ratings being more than one point apart. However, discrepancies of 11% for the cross-tab of scores assigned on the quality of claims were still acceptable, though with only 1 per cent excess of the 10 per cent threshold level as suggested by White (1984).

4.2.4 Interpretation

Three different statistical approaches were used to estimate the reliability of the obtained data (rater scores) of the present study. The results of which were triangulated to provide a ground for concluding whether the rater scores were reliable or not.

For correlational analyses, all tests were significant at the .01 level. Although the computed values of inter-rater reliability of both rating sessions appear to fall slightly short of the expected highs of .80s or .90s, the correlations between the rater scores and the resolved scores were much enhanced as a result of the negotiation

of discrepant scores between raters. To interpret the correlations, the values of r were squared to calculate the degree of overlap of two sets of rater scores. An increase in r value also causes an increase in r squared. Based on the correlation information, it would be too soon to judge whether reliability of the ratings was achieved or not. This could be explained from two statistical perspectives.

First, the value of correlation relates in part with the number of raters. This means that with the number of raters increases, the correlation value should increase too. Using the Spearman-Brown Prophecy formula to determine the number of raters to be used to obtain ratings with a .80 reliability estimate, the number of raters should be increased by 1.4 times (add 2 raters) for the holistic ratings, and 1.88 times (add 2 raters) for the analytic ratings. In other words, to achieve a .80 reliability estimate of the rater scores, 5 and 3 raters should be used for the holistic and analytic ratings, respectively. From the practical standpoint of the present study, ironically, only three (holistic ratings) and two (analytic ratings) raters were able to participate in the operational scorings of the students' essays. As the correlational analysis was conducted post hoc based on the rater scores, it is not possible to know in advance what the level of reliability will turn out to be.

The other factor attributing to the computed level of correlation of the rater scores is that correlation estimate does not assume a percent of discrepant scores (set at 10 % for this research) given by a pair of raters in comparison. While the first explanation suggests a solution to achieve a better reliability level, it could not be realized as raters were not available that much at the beginning of the ratings. To justify the reliability of rater scores, further investigation using less stringent criteria ,cross-tabulation, is necessary.

As the correlation results (r and r squared values) may reflect somewhat consistencies in raters' assigned scores, One-Way ANOVA was applied to examine whether there was any statistical difference between the mean scores of raters; that is, if some raters tend to give higher or lower scores than other raters. To test the reliability of rater scores, no significant difference between the rater mean scores was expected, indicating that their scores were close or similar to each other, or the ratings were highly likely to overlap. Unlike the test result of the holistic ratings, the significance difference exists for the analytic session's rater claim scores, and for the whole session. Again, like correlation, such difference could be attributed to the ANOVA procedure per se that uses the mean scores for significance tests and, therefore, discrepancies of scores were not its assumption. Taking this into account, such difference could be considered negligible as 1) the amount of difference between the reported p values and the alpha level were quite small (.004 for the claim scores, and .028 for the total ratings), and 2) by comparing the actual mean scores and standard deviations of the two raters assigned on the quality of written claim, and of the overall ratings, the amount of difference appears to be very close to each other and, as the general rule applies, not more than one point apart. For the overall ratings, the mean difference was .47, and the standard deviation differed by .11. For the analytic ratings as a whole, the differences were .26 and .08 for *M* and *SD*, respectively.

By adopting a less stringent criterion, a satisfactory level of rater reliability could be achieved by using cross-tabulation to check the frequency of discrepant scores between rater pairs. Based on White (1984), a level of discrepancies having one more point apart is expected not to exceed 10% of the ratings. With 10 per cent of disagreement allowed, cross-tabs results showed that only about 6 to 7 per cent of both

ratings were discrepant, and, therefore, the ratings were satisfactorily reliable. Theoretically, the cross-tab results for analytic ratings were consistent with what has been suggested to be one of the advantages of analytic scoring as multiple scores given to each script tend to improve reliability (Hamp-Lyons, 1991b; Huot, 1996). Excepting the level of disagreement between raters in claim scores which appears to be a bit higher than the expected level, the scores on the qualities of written reasons and refutations to counter-arguments could be considered an excellent reading with discrepancies less than 5 per cent as suggested by White (1984).

To sum up, the implications of computed test statistics discussed above support the ground for concluding that the data (rater scores) were reliable and appropriate for further statistical-inferencing procedures for investigating the research question (Research question 1).

4.3 Hypothesis Testing

All statistical output produced by SPSS for making decisions whether to accept or reject the null hypotheses relevant to the first research question are presented in Tables 4.11, 4.12, and 4.13 according to the types of test procedures used to test the hypotheses in question. Table 4.11 was discussed in sections 4.3.3. Tables 4.12 and 4.13 were elaborated in sections 4.3.4, and 4.3.5, respectively. All the resulting statistical values are summarized in Table 4.14 and discussed.

4.3.1 Test Variables

To evaluate and compare the effect of the instruction on the subjects' performance in their production of an argumentative essay, a number of statistical significance analyses were done using the three types of data obtained from the holistic and analytic scoring sessions. According to the hypotheses, the data were 1) the resolved holistic scores assigned to both groups' final drafts; 2) the resolved analytic scores assigned to both groups' first drafts; and 3) the resolved analytic scores gained by both groups' final drafts (term papers). It should be noted that the analytic scores comprise of the three separate scores awarded on the scripts on the rhetorical qualities of written claims, reasons, and rebuttals to counter-arguments, and not on the quality of language use. Such three types of data were treated as dependent variables (resolved holistic and analytic scores), and the instructional treatments as the independent variables. All significance tests were done at the .05 alpha level.

4.3.2 Levene's Test for Equality of Means

Since one important requirement of t-test that assumes homogeneity of variance has to be maintained, Levene's Test for Equality of Variances was computed to test the null hypothesis that the variances of the dependent variables are equal. For all dependent variables across the subject groups, the significance/probability values ($\underline{p}$) for the Levene test were high and greater than .05, indicating that the two groups' variances were not significantly different ($\underline{p}$ <.05). With the assumption maintained, the results of the t-test for equal variances were used and reported for concluding whether to reject or accept the null hypotheses of the study.

4.3.3 Independent-Samples t-Test (Nondirectional 2-tailed) & Eta squared

The two-tailed t-procedure was applied to test whether the experimental and control groups were different from each other in their argumentative writing performance as measured by their holistic and analytic gain scores. The results of the two-tailed significance tests are displayed in Table 4.11.

Table 4.11
Differences Between Groups in Holistic and Analytic Gain Scores on First- and Final Drafts

	Levene's Test for Equality of Variances		t-test for Equality of Means (Equal variances assumed)					Eta squared (η^2)**
	Sig.	$\underline{t}$	df	Sig. (2-tailed)	95% Confidence Interval of the Difference Lower	Upper		
Holistic scores (final draft)	.954	*.333*	1.391	53	*.170*	-.252	1.395	*.035*
1ˢᵗ Draft								
Claim	2.797	*.100*	2.342*	53	*.023**	.073	.945	*.094*
Reason	.030	*.863*	1.785	53	*.080*	-.053	.919	*.057*
Refutation	.494	*.485*	2.535*	53	*.014**	.123	1.059	*.108*
Final draft								
Claim	3.115	*.083*	2.076*	53	*.043**	.017	.988	*.075*
Reason	.021	*.886*	.749	53	*.457*	-.333	.730	*.010*
Refutation	.035	*.853*	2.047*	53	*.046**	.011	1.087	*.073*

*significant at the .05 level whereby $\underline{t} > 2.01$, and $\underline{p} < .05$
**$(\eta^2 \geq .14)$ = large effect size; $(\eta^2 \geq .06)$ = medium effect size; $(\eta^2 \geq .01)$ = small effect size

For the difference in the holistic scores gained by both groups' final drafts, the reported probability ($\underline{p}$) value of .170 ($\underline{t} = 1.391$) suggests no significant difference in argumentative writing performance on the final drafts between the experimental and control groups. Also, the tests were not significant for the difference in the rhetorical quality of written reasons demonstrated in the first- and final drafts with the

p values equal to .080 (t = 1.785) and .457 (t = .749) for the former and the latter pieces of writing, respectively.

In contrast, the tests were significant for the differences between the two groups in the scores on the rhetorical qualities of written claims and refutations to counter-arguments expressed in their first- drafts with the p values of .023 (t = 2.342) and .014 (t = 2.535) for the former and latter rhetorical qualities, respectively. For the difference in such two rhetorical aspects demonstrated in the final drafts, the 2-tailed probability values of .043 (t = 2.076) and .046 (t = 2.047) suggest a significant difference between the two groups in the two rhetorical aspects of written claims and refutations to counter-arguments, respectively.

The last column of Table 4.12 labeled Eta squared (η^2) displays the values of the effect size statistic for an independent-samples t-test. Eta Squared (η^2) was computed to measure how much of the difference in the writing performance (the dependent variable) of the two groups was really related to the different instructional treatments distinct to each of the two groups (the independent variable). η^2 ranges from 0 to 1. The greater the η^2 value closer to 1, the higher the proportion of variance in the dependent variable (mean scores) can be accounted for by the independent variable (instructional treatments), and vice versa. As a general rule, η^2 values of .01, .06, and .14 represent small, medium, and large effect sizes, respectively.

For the difference in the holistic mean scores on the final drafts of the two groups, the reported η^2 value of .035 was marginal of the extreme small effect size of .01. This means that only a small proportion of the variance in the holistic mean scores can be attributed to the instruction factor. For the remaining cells of η^2 value, a relatively fairly large effect size (η^2 = .14 more or less) was reported for the analytic

mean scores of the first drafts on the qualities of claim and refutation to counter-argument (η^2 = .094, and .108, respectively). In contrast, a relatively medium effect size was reported for the analytic mean scores of the first drafts on the quality of reasons used (η^2 = .057), and of the final drafts on the qualities of claim (η^2 = .075) and refutation to counter-argument used (η^2 = .073). The last and lowest eta squared value of .01 was reported for the difference in the final essays' mean score on the quality of written reasons.

4.3.4 Independent-Samples t-Test (Directional 1-Tailed)

To investigate whether the scores gained by the experimental group were actually higher than those gained by the controls (see Tables 4.1 and 4.3), the directional t procedure was used and the results reported in Table 4.12.

Table 4.12
Summary of One-Tailed Test Statistics

	Directional one-tailed independent-samples t test (equal variances assumed)		
	df	computed t	critical t (df = 53)	1-tailed .05 rejection level
Holistic scores (final draft)	53	*1.391*	1.67	Retain H_0
Analytic scores (1ˢᵗ draft)				
Claim	53	*2.342**	1.67	Reject H_0
Reason	53	*1.785**	1.67	Reject H_0
Refutation	53	*2.535**	1.67	Reject H_0
Analytic scores (final draft)				
Claim	53	*2.076**	1.67	Reject H_0
Reason	53	*.749*	1.67	Retain H_0
Refutation	53	*2.047**	1.67	Reject H_0

*significant at the .05 level with t > critical t

As can be seen from the table, the computed t values were greater than the critical t-value for most of the test variables, meaning that the experimental group gained significantly higher scores in most variables tested except the gain scores on the formulation of reasons. However, such significantly higher gain scores were much more obvious for the first drafts than the final drafts, considering the amount of difference between the computed t and the critical t.

4.3.5 Matched-Pairs t-Test

To test the null hypothesis of no significant difference within each group between the mean scores of first drafts and final drafts of argumentative essays, a matched-samples t-test was used. The results were shown in Table 4.13.

Table 4.13
Within-Groups Differences in 1ˢᵗ and Final Drafts' Analytic Scores

	t	df	Sig. (2-tailed)	95% Confidence Interval of the Difference Lower	95% Confidence Interval of the Difference Upper	Eta squared (η^2)**
Experimental group						
Pair 1: Claim	-2.423	27	*.022**	-.33	-.03	*.28*
Pair 2: Reason	-3.576	27	*.001**	-.51	-.14	*.32*
Pair 3: Refutation	-2.121	27	*.043**	-.28	.00	*.14*
Control group						
Pair 1: Claim	-2.431	26	*.022**	-.34	-.03	*.18*
Pair 2: Reason	-5.701	26	*.000**	-.76	-.36	*.56*
Pair 3: Refutation	-2.431	26	*.022**	-.34	-.03	*.18*

*significant at the .05 level ($p < .05$)
**($\eta^2 \geq .14$) = large effect size; ($\eta^2 \geq .06$) = medium effect size; ($\eta^2 \geq .01$) = small effect size

The computed test statistics reveal a significant difference across all matched-pairs tests with the two-tailed probability levels lower than the alpha level of .05, and 95% confidence intervals containing no zero value. For the comparisons between the mean scores of the final and first drafts of the experimental group on the qualities of claims, reasons, refutations to counter-arguments demonstrated by the writings, the tests (sig. 2-tailed) were significant with p-values of .022, .001, and .043 for the matched tests of claims, reasons, and refutations, respectively. For the control group, the matched-pairs tests were also significant for all three aspects of writing performance measured.

The last column labeled repeated-measures eta squared (η^2) reports the effect size statistics that explain the variation in the within-groups difference between the gain scores on the final- and first drafts that was accounted for by the effect of instruction distinct to each subject group. Based on the criteria stated above, the computed η^2 values were high for all matched-pairs test variables ($\eta^2 > 14\%$ = large effect size), indicating that the within-groups difference in the mean scores between the two test variables (sets of scores on 1st and final drafts of each group) were substantially attributed to the instruction factor (the independent variable).

4.4 Discussion of Results

Since the main concern of the present study is to investigate the possible effects of using a genre-based approach to teach the students to write an argumentative essay, the first research question was then translated into a total of 19 hypotheses and the results of which were presented in Tables 4.11, 4.12, and 4.13 according to the appropriate test procedures used to test the hypotheses in question. For the purpose of

discussion as well as triangulation of the results, all necessary results were summarized in Table 4.14. Based on the performance characteristics reflected in hypotheses tested and the test procedures used to test such hypotheses, the results can be discussed with respect to the impact of the instruction on the students' argumentative writing performance in three aspects including 1) between-groups differences, 2) within-groups differences, and 3) the effect size of the instruction attributed to such difference in performance. For the following discussion, Table 4.14 is used throughout.

4.4.1 Between-Groups Differences

4.4.1.1 Between-groups difference in 1st and final drafts' analytic scores

As demonstrated by the two-tailed t-test values that determine the difference between the two groups of subjects in their 1st and final drafts writing performance, the experimental group appears to outperform the control group in the two argumentative writing aspects measured namely the formulation (the quality of content and organization) of claims and rebuttals to counter-arguments only, whereas in the other aspect, the formulation of reasons, both groups were not significantly different from each other. The same results apply to both 1st and final drafts of the subjects' argumentative essays.

By comparing the p values for the difference in the draft essay performance with those for the difference in the final-draft performance, it can be noticed, however, that although the experimental group did actually differ from the control group in the two writing aspects as described above, such differences were much more apparent in the 1st drafts than in the final drafts, given that the p values for

the difference in the 1st draft writing performance were much less than the .05

rejection level,

Table 4.14
Summary of Statistical Test Results

Test variables (dependent)	Null hypothesis (H$_0$)	Test procedure (.05 significance level)	Computed test statistics (p)	Decision
1. Final-draft holistic scores	H1: Holistic scores	Independent-samples t-test (Table 4.12)	.170, .035**	H1: Retain H$_0$
2. 1st draft analytic scores	H2.1: Claim H2.2: Reason H2.3: Rebuttal	Independent-samples t-test (Table 4.12)	.023, .094** .080, .057** .014, .108**	H2.1: Reject H$_0$ H2.2: Retain H$_0$ H2.3: Reject H$_0$
3. Final draft analytic scores	H3.1: Claim H3.2: Reason H3.3: Rebuttal	Independent-samples t-test (Table 4.12)	.043, .075** .457, .010** .046, .073**	H3.1: Reject H$_0$ H3.2: Retain H$_0$ H3.3: Reject H$_0$
4. 1st draft analytic scores	H4.1: Claim H4.2: Reason H4.3: Rebuttal	Independent-samples t-test (1-tailed test) (Table 4.13)	2.342* 1.785* 2.535* (critical t =1.67)	H4.1: Reject H$_0$ H4.2: Reject H$_0$ H4.3: Reject H$_0$
5. Final draft analytic scores	H5.1: Claim H5.2: Reason H5.3: Rebuttal	Independent-samples t-test (1-tailed test) (Table 4.13)	2.076* .749* 2.047* (critical t =1.67)	H5.1: Reject H$_0$ H5.2: Retain H$_0$ H5.3: Reject H$_0$
6. Experimental group's analytic mean scores (draft & final)	H6.1: Claim H6.2: Reason H6.3: Rebuttal	Dependent-samples t-test (Table 4.14)	.022, .28** .001, .32** .043, .14**	H6.1: Reject H$_0$ H6.2: Reject H$_0$ H6.3: Reject H$_0$
7. Control group's analytic scores (draft & final)	H7.1: Claim H7.2: Reason H7.3: Rebuttal	Dependent-samples t-test (Table 4.14)	.022, .18** .000, .56** .022, .18**	H7.1: Reject H$_0$ H7.2: Reject H$_0$ H7.3: Reject H$_0$

*Computed t-value
**(η^2 ≥ .14) = large effect size; (η^2 ≥ .06) = medium effect size; (η^2 ≥ .01) = small effect size

compared with the difference between the p values for the difference in the final-draft performance and the .05 level. For 1ˢᵗ drafts, the differences between the p values and the .05 alpha level were .027, and .036 for the qualities of written claims and rebuttals to counter-arguments, respectively. In comparison, for the final drafts, the amounts of differences were .007, and .004 for the written claims and rebuttals to counter-arguments, respectively. For the between-groups difference in the quality of written reasons shown in their 1ˢᵗ and final drafts, both groups were equivalent in such aspect of writing quality even though the difference was almost significant for the draft-essay performance as suggested by the p values of .08 (>.05 level), compared with an insignificant difference for the final-draft performance with p equal to .457 (>.05 level).

The results derived from the one-tailed tests seem to give support to the two-tailed findings described above. With the amount of difference between the significant t levels and the critical t of 1.67 across the three quality aspects measured for 1ˢᵗ drafts greater than that for the final drafts, the similar conclusions to the two-tailed test results can be reached that the experimental group outperforms the controls in the writings of claims, reasons, and rebuttals to counter-arguments in the first drafts, and only in the formulation of claims and rebuttals to counter-arguments in the final drafts.

Up to this point, it can be noticed that the resulting statistical values derived from the one-tailed and two-tailed t-tests for determining the subjects' performance on the quality of written reasons look somewhat contradictory to each other for making a clear-cut conclusion to the hypotheses in question. With the experimental group's mean score on the first drafts' quality of written reasons higher

than the controls', the two-tailed p value of .080 which is more than the rejection p level of .05 render the mean difference between the two groups insignificant, suggesting no meaningful between-groups difference in the quality of written reasons, whereas the computed one-tailed t value of 1.785 (> the critical t value of 1.67) leads to the conclusion that the mean score on the quality of written reasons received by the experimental group was significantly higher than that of the control group.

Such somewhat conflicting results can be resolved by considering how close the computed t were to the critical t of 1.67 and the p-values were to the rejection p-level of .05. The prudent choice is suggested to be conservative in interpreting the results by ignoring the result of one-tailed t-test since the value of one-tailed test is only slightly higher than the critical t by .11, compared with the amounts of difference of .67 (2.34-1.67) and .86 (2.535-1.67) for the significant tests on the qualities of claims and rebuttals to counter-arguments, respectively, and adopting the two-tailed test result in which the difference on the quality of reasons of .03 (.08-.05) is much more in the same range as those on the qualities of claims (.05-.023 = .027) and rebuttals to counter-arguments (.05-.014 = .036).

The conclusion, then, is that the significantly higher mean score on the quality of written reasons gained by the experimental group's first drafts as indicated by the one-tailed test result (t = 1.785) might not necessarily imply that the experimental students perform better than the controls in the quality of written reasons in a meaningful way, and that the two-tailed test result suggesting a barely insignificant difference in such writing quality aspect (p = .08) was adopted instead.

4.4.1.2 Between-groups difference in the final-draft holistic scores

Although the experimental group gained (insignificantly) higher holistic mean score than the control group for an overall impression of the scripts (*2.28 v. 2.00*) as indicated by the one-tailed test result ($\underline{t}$ < critical $\underline{t}$), both groups still did not significantly differ from each other in their final-draft writing performance as suggested by the two-tailed test results of no significant between-groups difference in the holistic gain scores awarded on the final drafts. This is consistent with the assumption made at the outset of the study that there will be no significant difference between both groups in the final-draft writing performance.

4.4.2 Within-Groups Difference Between First-and Final Draft Analytic Gain Scores

The results of all matched-paired tests indicate a significant change within each group in the gain scores from first to final drafts for all three rhetorical qualities measured. For the qualities of written claims and reasons, both groups show an equivalent improvement in performance as shown by the $\underline{p}$ values which were the same or very close to each other in the degree of change in the qualities of written claims and reasons. In contrast, for the quality of written rebuttals to counter-arguments, both groups were quite different from each other. For the experimental group, an almost significant $\underline{p}$ value of .043 suggests a relatively less increase in the gain scores on the qualities of written rebuttals to counter-arguments of their final drafts, compared with the control group's $\underline{p}$ value of .022. This might probably signify that the control group's drafts have undergone much more improvement than the experimental groups' drafts as indicated by the larger amount of difference

between the scores gained by the 1st and final drafts in comparison to the experimental group although such significant development as indicated by the matched-samples tests was not adequate for making the control group outperform the experimental group in such rhetorical aspects of argumentative writing as shown by the results of independent-samples t-tests mentioned earlier, taking into account the fact that the mean scores on rebuttals to counter-arguments for the experimental group's 1st and final drafts were higher than those of the control group.

4.4.3 Effect Size Estimates

Since the matched-pairs test values tell only whether the student writers did improve or not following the instruction period of 12 weeks and nothing more than that, the effect size estimates or the so-called eta squared index were computed to approximate how much of such improvement in the writing performance could be accounted for by the instruction.

As described above, the effect sizes were estimated for all significant results of the two-tailed nondirectional, and matched t-tests. For the insignificant results, the eta squared were also computed although that might be redundant since such insignificant differences usually could be interpreted as the equivalent impact of the instructional treatments and were nearly always indicated by a very low value of eta squared. As shown in Table 4.14, the reported effect size estimates were varied across the test variables for each significance test.

For the matched-pairs tests, the computed eta squared values which were all more than 14% suggest a strong association between the instruction and the change in gain scores from 1st to final drafts of argument essays. This means that 14% or more

of the variance in the final-draft gain scores can be accounted for by the first-draft gain scores. In other words, the higher the level of significance as a result of the tests, the greater the effect size of the instruction. Such large effect sizes took place in both the control and experimental groups. For the gain scores on the qualities of written claims and rebuttals to counter-arguments, both classes were not different from each other in the developmental level of performance as influenced by the instruction distinct to each of the two classes. In contrast, the instruction could probably contribute much to the improvement in the formulation of reasons in the students' essays considering a very significant probability value of .001 associated with the eta squared of 32% for the experimental group, and that of .000 with the effect size of 56% for the control groups.

For the independent-sample tests, the effect size estimates appear to be large for the significant difference between the two groups in their analytic gain scores on the first rather than on the final drafts. For both types of drafts, the medium to large effect size was estimated for the qualities of written claims, and rebuttals to counter-arguments. However, for the quality of reasons, the impact of instructional treatments was notably different. This is shown by a medium effect size estimate of 5.7% for the experimental group, and a small effect size estimate of 1% for the control group.

4.5 Summary

To answer the research question, it can be concluded that the two groups of subjects participated in this instructional study did actually differ in argumentative writing performance as a result of exposure to two different types of argumentative writing instruction. For the experimental group, the use of genre-based approach to

the teaching of argumentative writing has proved to be effective as demonstrated by the statistically significant higher gain scores given to the experimental students' 1st and final drafts on the two rhetorical quality aspects of written argumentation, which are the development and organization of claims and refutations to counter-arguments, in comparison to the control group. However, such difference in performance was much more significant as shown by the 1st than the final drafts. For the quality of written reasons, both groups were not different from each other as suggested by the insignificant difference in their holistic gain scores on the quality of written reasons reflected in their first as well as final drafts of argumentative essays.

With respect to the impact of the instruction as shown by the computed eta squared index, it can be said that the effectiveness of the genre-based instructional approach could be realized as early as in the first drafts as suggested by the significantly higher scores gained by the experimental class's first drafts on the qualities of written claims and refutations to counter-arguments, in comparison with the control group. Nevertheless, it cannot be said that such effects have downplayed the likely merits of process writing practiced by the teachers of both classes since at the end of the course, both groups did improve in their developed performance as indicated by a dramatic change in gain scores as measured by a statistically significant difference in the gain scores between the final- and first drafts, and as evident in an insignificant difference between both groups in the holistic scores awarded on their final essays. Although the experimental group can be considered to still perform better than the control groups as demonstrated by their final drafts, the two-tailed probability values imply a very barely significant difference in the gain scores on

written claims and rebuttals to counter-arguments only. For the quality of written reasons, the difference between both groups' gain scores was found to be insignificant.

4.6 Questionnaire Responses

A retrospective questionnaire was administered at the end of the 12-week instruction period to the experimental group and responses returned from 19 respondents out of a total of 28 (a response rate of 65%). Important findings were summarized and discussed below. All summarized results are shown in Appendix C.

On the whole, the student respondents seem to be extrinsically motivated to learn English as they perceive that acquisition of English language skills such as reading and writing is important and can give them a competitive advantage over rivals in their specialized fields of studies and future career development. However, many of them find it difficult to write in English as they admit they often have trouble using correct grammar, sentence structure, and appropriate vocabulary to express their intended meanings. In addition, the school genres and writing assignments they have been acquainted with before attending English 5 were of free writing types such as summarizing and narration with the length of such writings of only about 1 or 2 paragraphs or not more than 1 or 2 pages at the most.

In writing an English argumentative essay in English 5, the student respondents seem to provide an essentially positive reaction to the instruction they were exposed to. This is evident in a positive feedback on how and in what ways they find the instruction useful in writing an argumentative essay on their chosen topics. In other words, the findings derived from the way the students perceived of the

instruction seem to reinforce the results of the hypotheses tested described in the previous section.

Based on the responses to the second set of question items, the instruction is perceived by the majority of students to be facilitative in helping them to develop and arrange their ideas coherently in their essays despite the fact that the development and organization of contents are considered by many of them to be the two problematic areas in writing in addition to the use of correct grammar and vocabulary. Since the students are encouraged to write several drafts and revise several times according to the teaching of process writing, the responses suggest that the instructional effect is obvious in drawing the learners to pay more attention to the contents or the rhetorical aspects of development and organization of ideas in the writing of their first drafts of argument essays, whereas the student writers focus relatively more on form (editing) in the stage of writing up their final drafts. As shown by the resulting test statistics described earlier, the focus on developing and revising the essay contents in the early drafts of the students' argumentative essays could, to a greater extent, contribute to the significant difference in the analytic gain scores on the first drafts between both groups of students.

Responses to the last part of the questionnaire represent the written feedback from the students on the usefulness and personal opinions they had about the conduct of the course. Since it seems likely that the students felt reluctant to express their opinions in such a way that criticizes the teaching authority, such written responses were obtained from only 5 student respondents. Regardless of the rate of response, the students' voices have proved to be very interesting.

With respect to the lessons and learning activities with particular reference to the prewriting stage, most of them similarly assert that they benefit most from the writing models and analysis of language form and function used in the models in helping them a lot in dealing with the researched information for their written topics and in using that information to write in such a way that conforms to the model, in sensitizing them to the writing convention of English argumentation, and for some of them, in letting them know at the beginning of the instruction what they were expected to do and how they were going to achieve that. One student has commented that:

> I think I like the way the teacher has taught me because it is quite comprehensible. At least I can acknowledge what my writing is expected to be…..that makes me feel I've got something to start with such as the models and materials provided in the textbook that I can always refer to when writing. In addition, I'm so surprised that I can finish my draft on my own and submit it to my teacher at the end of the 6th week which is much earlier than my friends studying in other sections.

Through informal conversation with the researcher, another senior transportation engineering student has asserted that:

> This is the second time I take this course. The first time I absolutely failed. I thought the problem was my very poor and non-standard English though I knew very well the subject matter of the topic I wanted to write. This is the reason why I kept postponing the course until my fourth year at the university. In this term, I think I could do better. My strategy is to follow and adapt the models in one way or another. Though this may look like copying…..by the way, the example patterns could have helped me a lot. I think the textbook should contain more example model essays.

CHAPTER 5

DISCUSSION AND CONCLUSION

This chapter consists of four sections. The first section discusses the results of the study in terms of the realized effects of explicit genre instruction on the student writing performance. The second section proposes pedagogical implications of the findings, followed by discussions on limitations and delimitations of the study, and directions for future research in section three and four, respectively.

5.1 Discussion and Conclusion

This quasi-experimental study using control group has taken a step in investigating the effects of using a genre-based approach in teaching EFL students to write argumentative essays. The results of this study clearly indicate that a teaching approach focusing on rhetorical organization can be successful in an EAP/EST teaching situation with less able or novice student writers.

Although the experimental and control instructional treatments were identical that both groups were taught using the same process approach, it is the genre-based approach integrated into prewriting instruction given to the experimental class that makes the experimental treatment different from that given to the comparison group. Thus, differences in outcomes between treatments can be attributed to the difference in pre-writing instruction, other things being equal. Within the scope of such distinction in terms of prewriting instruction, some possible key explanations can

account for the better performance of the experimental group in comparison to its counterpart.

5.1.1 Effect of Explicit Teaching of Genre in Prewriting

The research outcome that the experimental class can perform better than the control group as demonstrated by the analytic gain scores awarded on their draft essays' quality of development and organization of ideas or contents may suggest some sort of the instructional benefits that account for such superior performance. According to the assumption made at the outset of the study claiming the likely effectiveness of genre-based approach in prewriting on learning outcome as would be demonstrated by the experimental subjects' first-draft writing performance, the results then seem to affirm such projection. In addition, such learning gains can be achieved in a relatively short period of time as the experimental students can finish and submit the whole essay draft at the 6th or 7th week of the 12-week instruction period of the course in comparison to the control group.

One sensible explanation for the better performance is that explicit teaching of genre knowledge in prewriting can help student writers overcome or at least mitigate their difficulties when planning and attempting to write their argumentative essays. As pointed out in the literature, the difficulties faced by EFL students when asked to produce an academic piece of writing are often due to an inadequate understanding of how texts are organized to convey its communicative purposes (Hyland, 1990), frequently resulting in a production of incohesive and incoherent texts. Also, L2 student writers often tend to produce a narrative type of writing or engage in what Bereiter and Scardamalia (1987) called 'a knowledge-telling process,'

(as opposed to 'a knowledge-transforming process' engaged in by more able writers) or mainly reproducing information they know and remember in a form not valued by college/university academic writing.

By acknowledging that L2 learners may have only a limited competence in using alien discourse forms particularly of the genre they are expected to produce and that accessibility to such discourse forms is hardly possible in the situation where English is not used widely outside the classroom, the case is made for learners to be taught directly explicit knowledge of text structure to enable them to shape their work to the conventions of the target genre. As suggested by writing scholars, researchers, and teachers, this calls for a more interventionist (explicit) approach in the process writing classroom in teaching learners what they need to know (Hillocks, 1986; Walberg, 1990; Larsen-Freeman and Long, 1991; Rosenshine and Guenther, 1992). Writing researchers like Hillocks and Walberg have found that teachers "promote learning, *first by making explicit their tacit knowledge or by modeling their strategies* for students in authentic activity" (cited in Williams and Colomb, 1993, p. 257, italics added). This teaching practice is referred to as "explicit scaffolding" by Rosenshine and Guenther (1992).

An important aspect of using genre-based approach in the context of the present study is to have students work with the whole text at the beginning, requiring that before attempting to write on their own, they need to have been exposed to the genre by reading, analyzing, and discussing examples of it. In this study, the approach used in prewriting involves three main phrases: modeling, joint text construction, and independent construction of text. By discussing the characteristic text features through the presentation of model texts, students begin to understand how and why

texts are organized and interact with its communicative purposes in certain ways. The genre knowledge that comes from discussing models also helps students to understand their learning goals, what they are expected to perform, and provide them with guidelines for evaluating their own writing as well as heuristics for planning and organizing their materials necessary for writing up their first drafts.

In retrospect, explicitly teaching the structures and grammatical features of text through joint text construction using an overhead projector might be viewed to be one beneficial aspect of scaffolded classroom interaction as guided practice used by the experimental-class teacher. Through such an instructional technique, students were made aware of rhetorical strategies and language points necessary in the assignment genre as the teacher made think-aloud comments while planning and creating texts for the overhead. As Randi Reppen has suggested, this gave students "valuable insights into decisions made during text construction and provided them with opportunities and tools to talk about language" (1995, p.33). This kind of scaffolded interaction, Hawkins argues, provides "abundant opportunities for linguistic and cognitive modeling and development" (1988, p.128). Once students become familiarized with the characteristic text and language features of the task they are expected to perform, they increasingly progress to the point where they are able to perform a task themselves. In other words, they move toward their own performance of a task, first assisted through joint text construction, and then independently constructing text on their own.

Empirical research has shown learning gains through scaffolded classroom interactions, suggesting how language learners are able to reach higher levels of performance than they would have been on their own (e.g., Hawkins, 1988; Tharp and

Gallimore, 1991; Kowal and Swain, 1994; Donato, 2000; Ohta, 2000). Scaffolded classroom interaction, as Hawkins argues, is likely to be effective when the situation is both interactively and cognitively demanding for students to be capable of higher-level cognitive activities that is beyond their level of competence (1988, cited in Paltridge, 2001). This suggests that at least in the early phase of instruction, students are dependent on the teacher for input and explicit instruction.

To examine the effect of explicit instruction (explicit scaffolding) in prewriting, some parts of essays selected from both groups can be compared and contrasted in Figure 5.1. While the introductory and concession paragraphs were pointed out by the analytic gain scores to be the two highly contrastive areas of student writing, the introductory paragraphs drawn from each of both groups' students were selected from their drafts for discussion (Table 4.15). It should be noted that the excerpts were not typical characteristics of each group but should be considered as examples contrasting the likely instructional effects on both classes' first-draft performance. To focus on the students' ability to control rhetorical elements, the excerpts were corrected for some awkward grammatical mistakes and inappropriate vocabulary without change in meanings, contents, and organization.

In teaching writing as a process, teachers usually attend to students' ability in developing and organizing contents and ideas in their preliminary drafts before turning to focus on accuracy or language use in the later stages of the process. As the students participated in this study were required to produce an argumentative text of a problem-solution pattern, it is clear to any teacher reader/rater that the excerpts drawn from both groups are highly contrastive at both ideational and organizational levels. For the introductory paragraphs, the differing points can be considered in terms of

how ideas are presented and organized around the purpose or the writer's position, the situation where the problem occurs, and the suggested solution. Since this study concerns mainly with measuring the effect of instruction in prewriting, examples of student writings should be analyzed in terms of how ideas were developed and organized in their first drafts as shown in Figure 5.1.

Figure 5.1 Contrasting the Introductory Paragraphs

An introductory paragraph written by one control student:

> In all nations, intermodal transportation is important because it can reduce varied expenses specially transportation cost and transportation time. It can help foreign nation commerce grow and has efficiency. But in Thailand intermodal transportation is not popular because the commerce uses direct transportation for example trains, trucks, airplanes and shipping. But this transportation is old transportation. It has many problems to commerce such as high transportation cost when airplanes are used to carry. Furthermore, trucks cannot carry a lot of products. These are examples of why commerce grows slowly and is inefficient. Thailand is a country who have many resources for intermodal transportation including trucks, trains, ships, and airplanes. They can apply to many methods of intermodal transportation such as piggyback, shuttle trains, etc. All methods have different advantages. Some method can help about goods transport. But some methods help about traffic in city and downtown. All methods can help reduce problem in Thailand for more efficiency in commerce. Then intermodal transportation should be used in Thailand for more efficiency in commerce.

An introductory paragraph written by one experimental student:

> Drying of longan is becoming a very interesting business because of long storage and increased price of product. Drying longan can be an excellent alternative for longan farmers who own longan garden with too many longans during the seasons. Also, if farmers can add a small factory for drying longan, they can get dried longan meat, which is another product that has export potential. If longan growers want to invest in small factory, they may consider buying drying machine from the quality producers with good after-sale service. It is also recommended that the drying machine should be tray- or cabinet dryers. The cabinet dryers have heated by means of a heater at the entrance and forced through the stack of trays, so it can dry longan with quality. However, drying longan using cabinet dryer may cause flavour changes and decrease shelf life of the product because direct combustion of natural gas used in heating and drying longan in cabinet dryer can come in direct contact with the food products. Although cabinet dryer is more variable in controlling the quality of product, longan grower should use cabinet dryer for drying longan in small factory because it is suitable for small-scale drying, has high quality, and is easy for maintenance.

By not taking into account the students' problem of language use, it is obvious that both paragraphs were very much different from each other with respect to the rhetorical quality of generated ideas and how ideas were organized in the paragraphs. For instance, on the topic of 'intermodal transportation,' the control-group student may not be able to successfully and sufficiently introduce the topic to the reader, leaving much to the reader's own interpretation or guesswork whether the problem really exists and the proposed solution is practical and makes sense. In comparison, the paragraph written by the experimental student seems to be easier for readers to read and understand as the writer has attempted to situate text in a specifically defined rhetorical context including target audience (longan growers), problematic situation (preserving excess longans for out-of-season sales), and proposed solution (using cabinet dryers for drying longans). The better performance as demonstrated by this experimental student's writing of introduction may reflect the usefulness of explicit teaching of rhetorical form and function in addition to that of abstract teaching of argumentation concepts. For some writers, the effect of teaching focusing on the abstract principles alone may result in the writing that lacks clarity and focus. However, it should be noted that the control student's paragraph is not representative of the control class. Rather, it highlights a likely case when teaching only necessary concepts for students to write up their early drafts may not be adequate at least for some student writers.

These examples of student writing thus reflect a difference in learning gains that can be accounted for by the difference in prewriting instruction. As indicated by the calculated effect size estimates (eta squared index), a significant difference between both groups' gain scores on first draft essays is to a considerable extent

influenced by the effect of instruction particularly on the rhetorical qualities of written claims and refutations to counter-arguments, though the impact of instruction on the quality of written reasons is found to be very small. This may suggest that the development and organization of claim or thesis and of refutation to counter-argument may be the two important aspects of argumentation skills that without explicit genre knowledge, the learning outcome of both groups of students might not be significantly different from each other.

5.1.2 Possible Pitfalls of Genre-Based Instruction

While the genre-based instruction has been proved to have powerful pedagogic potential as evident in this study and elsewhere, concern was expressed about its rigidity and prescriptiveness (e.g., Widdowson, 1983; Swales, 1990; Freedman, 1993a; Reppen, 1995; Kay and Dudley-Evans, 1998; Bhatia, 1999; Flowerdew, 2000; Swales, 2000b). Work on genre analysis and the use of its findings of typical textualizations, as Widdowson has commented, might lead practitioners "to suppose that form-function correlations are fixed and can be learnt as formulae" (1983, p. 102).

A common concern among these composition researchers centers around the tendency that genre approach may limit student expression through the use of model texts and so may turn genre as a teaching tool into a matter of applying formulaic prescription of how text should be structured. Such a possibility may give rise to what Swales terms 'formulaic parroting' (1990, p. 16), leading to the situations in which students are instructed to manipulate certain features (Reppen, 1995), to further replicate genre in similar rhetorical contexts (Bhatia, 1999), and for some

students, to consider it as a rigid and prescriptive model for them to emulate blindly (Swales, 1990; Kay and Dudley-Evans, 1998).

In this research, based on raters' shared experience, it is fortunately enough to have observed that such a case rarely exists and should be considered exceptional, and in the cases where such concern was evident, it may be quite acceptable. A possible explanation for this is probably one advantage of contextualizing a text before its presentation by locating the genre in rhetorical situation through discussion of purpose, audience, and occasion. For instance, a student writer as an environmental engineer would need to consider a specifically defined rhetorical context of his argumentative writing including a specific circumstance in which a specific problem occurs (waste water from a paper mill factory) and a way of solving the problem as well as its competing alternative (oxidation pond versus activated sludge).

For a group of learners with relatively more advanced English proficiency, such instruction on metaknowledge of argumentation and elements of arguments as stipulated in the writing task may suffice in facilitating them to compose their preliminary drafts. For lower-level learners, such implicit knowledge of genre may not be adequate. In the latter case, explicit teaching of genre as discussed above provides examples of the target genre and appropriate language embedded in the genre that these students need in writing up their drafts. However, it is likely that when students tend to rely heavily on the model texts, their writings may appear to be somewhat arid and monotonous, and, in some cases, may result in the overgeneralization and overuse of language features presented in the example texts. An excerpt of one student's writing shown in Figure 5.2 illustrates an example case of prescriptivism of explicit modeling.

As can be seen from the excerpts, the student seems to overuse and sometimes misuse the language pattern provided in the model text with only change in key vocabulary (words underlined) representing the arguments supporting the use of biodiesel in buses in Thailand. Although the writing is well structured and organized and contains no serious flaw in grammatical accuracy, the overuse and misuse of some lexical phrases may affect the overall impression of the writing when the essay was rated holistically.

Figure 5.2

Likely Prescriptive Effects of Explicit Modeling of Genre

Body paragraph 1:

The first reason why biodiesel should be used in the buses in Thailand is that biodiesel can be used to reduce <u>air pollution</u> whereas petroleum diesel cannot.………However, this is no such evidence to show that petroleum diesel is also proved to be useful. In short, we can say that biodiesel is better than petroleum diesel in that biodiesel can be used to <u>reduce air pollution</u> better than petroleum diesel in the buses in Thailand.

Body paragraph 2:

The second reason why biodiesel should be used in the buses in Thailand is that biodiesel can be used to <u>reduce greenhouse gas emission</u> while petroleum diesel cannot……….. However, there is no research done so far to prove that petroleum diesel can do same thing. In Thailand, a study by Jareanrung (2003, p.345) shows that biodiesel can <u>reduce greenhouse gas emission</u>. However, there is no such evidence to show that petroleum diesel is also proved to be useful. In short, it is no doubt that biodiesel is better than petroleum diesel in that biodiesel can be used to <u>reduce greenhouse gas emission </u>better than petroleum diesel in the buses in Thailand.

Body paragraph 3:

The final reason why biodiesel should be used in the buses in Thailand is that biodiesel can be used to <u>save more energy</u> whereas petroleum diesel cannot……………..However, there is no research done so far to prove that petroleum diesel can do same thing. In Thailand, a study by Jareanrung (2003, p.126) shows that biodiesel can <u>reduce fuel consumption similar to petroleum diesel</u>. However, there is no such evidence to show that petroleum diesel is also proved to be useful. In short, we can say that biodiesel is better than petroleum diesel in that biodiesel can be used to <u>save more energy in vehicle in Thailand</u>.

For inexperienced writers, the problems of excessive and inappropriate replication require an appropriate pedagogical remedy. In fact, the problem of rigidity and prescriptiveness may not be entirely attributed to genre as a teaching tool per se but rather to inadequate knowledge of lexico-grammatical phrases necessary to be used interchangeably to convey the intended meanings demanded by textual features of the assignment genre. In many cases, it might not be easy for students to avoid such excessive repetition of key lexical phrases presented in the example texts because of lack of knowledge and skills of using language to talk about text. Students need to understand that the model texts are just possible patterns that make up the genre, not set patterns of form. The remedial action would be to expose students to some sort of language input such as a variety of flexible language patterns that can be used to achieve the intended meanings demanded by the genre. Doing so can provide students with more language options that they can use to organize and express their ideas more effectively. (See 5.2.2 for further teaching implications)

5.1.3 Affective Effect

Another notable outcome of explicit genre instruction may also be affective, giving students confidence when approaching writing tasks. Based on the questionnaire responses (See 4.4 and Appendix C), learning and writing experience student respondents have in common seems to reveal perceived facilitative benefits of the lessons, classroom activities, and materials, possibly reflecting a reorientation in their attitude toward EFL academic writing.

Following instruction, the subjects as novice writers reported a greater sense of ease when writing in English, an activity that has been previously quite stressful for

them. EFL students often report that they find it difficult to write in English and admit they dislike English composition. Such writing difficulty is frequently a matter of not being able to write because of a lack of knowledge of the informational and organizational demands of English rhetorical patterns and audience expectations. Once the rhetorical demand of the genre is explicitly made clear to students, their increased understanding of text features then makes their reading and prewriting purposeful, reducing the amount of time required before starting to write. And once started to write, students can write with much greater ease, as they are able to formulate their ideas more efficiently, concentrating on combining the elements effectively in terms of both achieving their intended communicative goals and conforming to the rhetorical conventions.

This positive affective outcome pointed out in this study then provides another piece of evidence affirming a body of research which has asserted L2 learners' positive development in attitude toward writing in English as a result of the beneficial effects of genre-based instruction (e.g., Reppen, 1995; Henry and Roseberry, 1998; Kay and Dudley-Evans, 1998; Hyon, 2001; Yoshimura, 2002). Kay and Dudley-Evans (1998) ran a multicultural workshop participated by 48 language teachers working in EFL/ESL environments, asking them what they thought of the notion of genre and its applications in the writing classroom. While the danger of genre-based approach was pointed out that it might become too prescriptive, many teachers considered the approach especially suitable for beginner and intermediate students, as it enabled student writers "to produce a text that serves its intended purpose" (p. 310). They further maintained that for learners with lower levels of English proficiency and literacy, offering them models was useful as they gave learners confidence and

security in approaching the writing task, liberating them from their own fears of writing and providing a means whereby they could analyze the effectiveness of their own writing and that of peers.

5.1.4 Conclusion

The instructional treatment used in this study provides a context of how genre-based instruction or explicit teaching of genre can be successful in the course of teaching writing as a process. The focus on the better performance as a result of the inclusion of explicit instruction in prewriting does not mean to downplay the potential effect of process writing instruction since at the end of the course both groups' final-draft writing performance was not significantly different from each other.

This research can be concluded that:

1) The experimental group receiving explicit teaching of argumentative genre can outperform the control group in their writing performance of first drafts on the quality of claims and refutations to counter-arguments but not on the quality of reasons;

2) Through the teaching and learning of process writing, both groups of students are not different from each other in their final-draft writing performance as indicated by the holistic gain scores;

3) Although the experimental students' final drafts gain significantly higher scores on the quality of written claims and rebuttals to counter-arguments than the controls as indicated by the analytic scores, such gains are barely significant and could be viewed as negligible; and

4) The insignificant difference between both groups of subjects in gain scores on the quality of written reasons may suggest that both classes may have experienced less difficulty in the development and organization of reasons used in their writing, in comparison with the other two aspects of argumentation evaluated.

The findings thus reflect:

1) genre-centered writing instruction could complement rather than contravene writing process instruction;

2) explicit genre knowledge helps students increase approximation to control the genre in less time, suggesting even greater gains over the extended period of instruction;

3) the beneficial effects of process writing instruction are measurable after a much longer period of instruction; and

4) teaching abstract knowledge of argumentation combined with explicit instruction of genre can better facilitate learning than teaching the concepts of argument alone.

5.2 Pedagogical Implications

The important findings of the present study that the experimental group can demonstrate significant gains in abilities to write a well-developed and organized argumentative essay as a result of prewriting instruction suggest two important aspects of instruction: 1) rhetorical and 2) language.

5.2.1 Rhetorical Focus

As shown by the findings of this research, an integration of genre-based instruction into a process writing teaching approach in which students plan, draft, revise, and edit their work is desirable and can provide an optimal learning condition for inexperienced and unskilled EFL writers.

EFL as opposed to L1 student writers generally lack implicit knowledge of rhetorical plans, organizational logic, and genre form, causing them difficulty in planning and starting to write. Teaching them genre analysis in prewriting through reading, discussing, and analyzing model texts from the beginning then is important in sensitizing them to the rhetorical patterns that are usually revisited in an English-speaking culture and a means of making meanings specifically stipulated by the writing task. Although explication of genre in this way has been concerned by some teaching professionals that the approach might lead students to expect to be told how a text should be structured and to replicate the model in a similar rhetorical context of his or her own writing, contextualizing genre in rhetorical situation involving audience expectation, occasion, and purpose can keep such undesirable possibility to a minimum.

While it is also possible to work from the student's text toward the genre through the teaching of composing process skills, this may require greater amount of time and efforts on the parts of both teacher and students. For students with lower English proficiency and literacy, this could mean too many drafts and correction according to teacher feedback. For a group of relatively more able EFL writers, on the other hand, instruction on abstract concepts of argumentative writing may be suffice in helping them compose their essays. Since a classroom usually consists of both types

of learners and that learners cannot be expected to know how to write a clear, focused, well organized and coherent text, making implicit knowledge explicit through genre-based instruction in prewriting could give EFL novice writers a level playing field in which they can know teachers' expectations and how to perform the task efficiently and effectively. As Bandura (1977) has noted:

> Learning would be exceedingly laborious, not to mention hazardous, if people had to rely solely on the effects of their own actions to inform them what to do. Fortunately, most human behavior is learned observationally through modeling: from observing others, one forms an idea of how new behaviors are performed, and on later occasions this coded information serves as a guide for action. Because people can learn from example what to do, at least in approximate form, before performing any behavior, they are spared needless errors (p. 22).

5.2.2 Language Focus

While students need to be familiarized to the rhetorical structure and organization of the genre, they also need to be made aware of the variety of key lexical phrases which are prevalently representative of the prototypical structures (Nattinger and DeCarrico, 1992). This is significantly necessary for helping students avoid the overuse or misuse of the key lexical phrases explicated in the model essays which could make their writing look arid and monotonous to readers who usually are their teachers and evaluators of their essays. Therefore, L2 writers need to be exposed to instruction and exercises on a greater variety of appropriate register for the rhetorical functions of argumentative writing. They can be given or asked to provide alternatives and encouraged to discuss with teachers or peers the contexts of their own writing.

Flowerdew (2000) suggests instruction on two aspects of language points that frequently co-occur with the problem-solution pattern of argumentative writing:

the lexical phrases used for textual organization and those appropriate for achieving the rhetorical functions of texts in contexts. In the former case, a number of lexical phrases typically used for discourse organization and expressions of causal relationships and making deductions and concessions, for instance, can be presented for the students to use interchangeably in their writing. In the latter case, although students are already aware of such phrases and use them in their writing, they often are not presented in an acceptable form or not appropriate for the contexts. For example, some students tend to overuse or misuse expressions when the context requires a more mitigating expression using a modal verb. For the statements that forward claims (theses/proposed solutions) and that acknowledge and respond to possible counter-arguments which are usually written in a way that suggests an attitude of tentativeness or modesty on the part of the author (Weissberg and Buker, 1990), it is conventional to sound more cautious rather than too sure of the benefits, either practical and theoretical, of the propositions (solution to the problem), given that there also exists competing solutions and benefits.

By introducing to the student writers the aforementioned lexico-grammatical features and their variations as metadiscourse (e.g., Vande Kopple, 1985; Crismore, Markkonen, and Steffensen, 1993; Intaraprawat and Steffensen, 1995; Hyland, 2000) that student writers can flexibly make use of them in achieving the rhetorical purposes of their texts, the quality of persuasiveness of their writing could be much enhanced, and would result in considerably improved writing styles and fluency.

5. 3 Limitations and Delimitations

The findings of this research should be interpreted with caution due any quasi-experiment, taking into account the following limitations embedded in the research components.

1. The number of raters used in the present study may fall slightly short of the optimal number of raters suggested by relevant empirical studies. While research has consistently shown high scorer reliability when four raters were used, the present study used only three raters for holistic scoring and only two of them for analytic scoring. However, a number of appropriate statistical procedures were used to warrant the reliability of rater consistencies in their assigned scores, and based on White's (1984) criteria, an adequately high level of inter-rater reliability was achieved.

2. As a quasi-experiment undertaken in formal educational settings, the subjects were not randomized and assigned to conditions. While it is possible to assume beforehand that the subjects of each group were of varied levels of language proficiency and, in particular, of different writing competence, the written responses to the argumentative writing prompt administered before the instruction began suggest the students' limited range of argumentative writing capabilities. This implies that if they were assigned to write an argumentative essay, it is highly likely that they would construct their own writing patterns of argumentation that may depart in some ways from English rhetorical conventions. In addition, a large sample size of nearly 30 students for each group could, to a certain extent, delimit the extent to which the present study's findings can be generalized. Statistically, a sample size of 30 or more can represent a true range of characteristics of the

population under investigation. As a matter of fact, the university registration system allows the students to select the classes as they see fit, and as such, the way the students are put into groups could be viewed as a random selection.

5.4 Directions for Future Research

To pursue the same line of research similar to the present study, modifications to the research components are recommended for further studies.

1. For further research, sample size and number of raters should be estimated in advance for the desired results to be achieved. To calculate the optimal number of raters needed to achieve high scorer reliability, the Spearman-Brown Prophecy formula can be used. Alternatively, given that the mean and standard deviation of the population were known, the optimal sample size can be estimated to achieve the desirable outcome of the significance test.

2. If feasible, a true experimental design may need to be undertaken for the results to be more accurate and generalisable to specific types of learners. Samples should be grouped according to their language proficiency particularly writing competence to determine if the instruction is effective across classified groups and treatment conditions. A discriminant analysis would be useful for this type of research.

3. Since randomization is rarely feasible in classroom research, replications of this type of instructional study may take place in different teaching institutions using different types of learners and different teachers.

4. Another research direction would be to investigate whether the use of different types of instructional input and exercises can be proved effective as early as in the

first draft essay as shown by the present study. While it has been shown that by drawing learners' attention to input which is comprehensible in the early classes through explicit teaching of form and function of texts, it remains to be seen whether giving students opportunities to engage in teacher-devised rhetorical consciousness-raising activities such as having them read and analyze more texts and distinguish between good and poor writing using a set of student essays of varied bands could enhance their comprehension and writing performance.

REFERENCES

Alderson, C. (1991). Dis-sporting life: Response to Alastair Pollitt's paper 'Giving students a sporting chance: Assessment by counting and by judging.' In C. Alderson & B. North (Eds.), *Language testing in the 1990s* (pp. 46-59). Hemel Hempstead, Hertfordshire: Phoenix ELT/Prentice-Hall UK.

Alderson, C. (1994). Bands and scores. In C. Alderson & B. North (Eds.), *Language testing in the 1990s* (pp. 71-86). Hemel Hempstead, Hertfordshire: Prentice Hall International.

Allaei, S. & Connor, U. (1991). Using performative assessment instruments with ESL student writers. In L. Hamp-Lyons (Ed.), *Assessing second language writing* (pp. 227-240) Norwood, NJ: Ablex.

Anthony, L. (2000). Implementing genre analysis in a foreign language classroom. *TESOL Matters, 10* (4), 18.

Applebee, A., Langer, J., & Mullis, I. (1986). *The Writing report card: Writing achievement in American schools.* Princeton, NJ: Educational Testing Services.

Applebee, A., Langer, J., Mullis, I., Latham, A., & Gentile, C. (1994). *NAEP 1992 writing report card (Report 23-W01).* Washington, DC: National Center for Education Statistics.

Aristotle. "Rhetorica." (translated by W. Rhys Roberts). In M. Adler (Ed.), *Great books of the western world.* (Vol. 8, pp.587-675). Chicago: Encyclopaedia Britannica.

Bachman, L. & Palmer, A. (1996). *Language testing in practice.* Oxford: Oxford University Press.

Badger, R. & White, G. (2000). A Process-genre approach to teaching writing. *ELT Journal, 54 (2),* 153-60.

Bamberg, B. (1983). What makes a text coherent? *College Composition and Communication, 34,* 417-429.

Bandura, Albert. (1977). *Social learning theory.* Englewood Cliffs, NJ: Prentice Hall.

Bandura, Albert. (1986). *Social foundations of thought and action: A Social cognitive theory.* Englewood Cliffs, NJ: Prentice Hall.

Batteiger, R. (1994). *Writing and reading arguments: A Rhetoric and reader.* Needham Heights, MA: Allyn & Bacon.

Bereiter, C. & Scadamalia, M. (1987). *The Psychology of written composition.* Hillsdale, NJ: Lawrence Erlbaum.

Bereiter, C. & Scardamalia, M. (1982). From conversation to composition: The Role of instruction in a developmental process. In R. Glaser (Ed.), *Advances in instructional psychology* (Vol. 2, pp. 1-64). Hillsdale, NJ: Erlbaum.

Bernstein, B. (1990). *The Structuring of pedagogic discourse: Vol. IV—Class, codes, and control.* London: Routledge Kegan Paul.

Bhatia, V. (1993). *Analysing genre: Language use in professional settings.* London: Longman.

Bhatia, V. (1999). Integrating products, processes, purposes and participants in professional writing. In C. N. Candlin & K. Hyland (Eds.), *Writing: Texts, Processes and Practices.* London: Longman.

Bitzer, L. (1968). The Rhetorical situation. *Philosophy and Rhetoric, 1,* 1-14.

Bruner, J. (1986). *Acts of meaning.* Cambridge, MA: Harvard University Press.

Burns, A. (1990). Genre-based approaches to writing and beginning adult ESL learners. *Prospect,* 5 *(3),* 62-71.

Burns, A., Joyce, H., and Gollin, S. (1996). *I see what you mean: Using spoken discourse in the classroom: A Handbook for teachers.* Sydney: NCELTR.

Callaghan, M., Knapp, P., and Noble, G. (1993). Genre in practice. In B. Cope & M. Kalantzis (Eds.), *The Power of literacy: A Genre approach to teaching writing* (pp. 179-202). Pittsburgh: University of Pittsburgh Press.

Campbell, D. & Stanley, J. (1963). Experimental and quasi-experimental designs for research on teaching. In N. L. Gage (Ed.), *Handbook of Research on Teaching.* Chicago: Rand McNally.

Chang, Y. & Swales, J. (1999). Informal elements in English academic writing: Threats of opportunities for advanced non-native speakers. In C. Candlin & K. Hyland (Eds.), *Writing: Texts, processes, and practices* (pp. 145-167). London: Longman.

Chaya, W. (forthcoming). *Effects of the explicit metacognitive strategy training on EFL students' revision of their persuasive essays.* Unpublished doctoral dissertation, Suranaree University of Technology, Nakorn-radchasima, Thailand.

Cheedlamduan, B. (1995). *Reading and writing ability and responsibility in learning English through the instruction based on genre-based approach minicourses and the instruction based on the teacher's manual.* Unpublished M.Ed. thesis, Srinakharinwirot University, Bangkok.

Chinnawong, S. (2002). Needs and preferences of EAP science student writing: Teacher versus student perspectives. *Thai TESOL Bulletin,* 15 *(1),* 7-22.

Clark, I. (1994). *Writing about diversity: An Argument reader and guide.* Orlando, FL: Harcourt Brace.

Clarke, M. (1994). The Dysfunctions of the theory/practice discourse. *TESOL Quarterly, 28,* 9-26.

Coe, R. (1994). Teaching genre as process. In A. Freedman & P. Medway (Eds.), *Learning and teaching genre* (pp. 157-169). Portsmouth, NH: Boynton/Cook.

Cohen, A. (1994). *Assessing language ability in the classroom.* (2nd ed.). Boston, MA: Heinle and Heinle.

Connor, U. & Kaplan, R. (1987). *Writing across languages: Analysis of L2 text.* Reading, MA: Addison-Wesley.

Connor, U. & Lauer, J. (1985). Understanding persuasive essay writing: Linguistic/rhetorical approach. *Text, 5 (4),* 309-326.

Connor, U. & Lauer, J. (1988). Cross cultural variation in persuasive student writing. In A. Purves (Ed.), *Writing across languages and cultures: Issues in contrastive rhetoric* (Vol. 2, pp. 138-159). Newbury Park, CA: Sage.

Connor, U. (1987). Argumentative patterns in student essays: Cross-cultural differences. In U. Connor & R. Kaplan (Eds.), *Writing across languages: Analysis of L2 text* (pp. 57-71). Reading, MA: Addison-Wesley.

Connor, U. (1990). Linguistic/rhetorical measures for international persuasive student writing. *Research in the Teaching of English, 24 (1),* 67-87.

Connor, U. (1996). *Contrastive rhetoric: Cross-cultural aspects of second-language writing.* Cambridge: Cambridge University Press.

Cooper, C. & Odell, L. (Eds.). (1998). *Evaluating writing: The Role of teachers' knowledge about text, learning, and culture.* Urbana, IL: National Council of Teachers of English.

Cooper, L. (1998). What we know about genres, and how it can help us assign and evaluate writing. In C. Cooper & L. Odell (Eds.), *Evaluating writing* (pp. 23-52). Urbana, IL: National Council of Teachers of English.

Cope, B. & Kalantzis, M. (1993). *The Power of literacy: A Genre approach to teaching writing.* Pittsburgh: University of Pittsburgh Press.

Crammond, J. G. (1997). *An Analysis of argument structure in expert and student persuasive writing.* Unpublished doctoral dissertation, McGill University, Montreal, Canada.

Crismore, A., Markkonen, R., & Steffenson, M. (1993). Metadiscourse in persuasive writing: A study of texts written by American and Finnish university students. *Written Communication, 10,* 39-71.

Crowhurst, M. (1983a). *Persuasive writing at grades 5, 7, and 11: A Cognitive-development perspective.* Paper presented at the Annual Meeting of the American Educational Research Association, Montreal. (ERIC Document Reproduction Service No. ED 230977).

Crowhurst, M. (1983b). *A Developmental perspective on persuasive writing.* Paper presented at the Annual Meeting of the Conference on College Composition and Communication, Detroit, MI.

Crowhurst, M. (1987). Cohesion in argument and narration at three grade levels. *Research in the Teaching of English, 21,* 185-201.

Crowhurst, M. (1991). Interrelationships between reading and writing persuasive discourse. *Research in the Teaching of English, 4,* 314-35.

Delaruelle, S. (1997). *Access.* National Centre for English Language Study and Research (NCELTR), Macquarie University, Sydney.

Delia, J., Kline, S., & Burleson, B. (1979). The Development of persuasive communication strategies in kindergarteners through twelfth-graders. *Communication Monographs,* 46 *(4),* 241-156.

Donato, R. (2000). Sociocultural contributions to understanding the foreign and second language classroom. In J. P. Lantolf (Ed.), *Sociocultural theory and second language learning.* Oxford: Oxford University Press.

Dudley-Evans, T. (1985). *Writing laboratory reports.* Melbourne: Thomas Nelson.

Dudley-Evans, T. (2001). The Teaching of the academic essay: Is a genre approach possible? In A. Johns (Ed.), *Genre in the classroom: Multiple perspectives* (pp. 225-235). Mahwah, NJ: Lawrence Erlbaum.

Durst, R., Laine, C., Schultz, M., and Vilter, W. (1990). Appealing texts: The Persuasive writing of high school students. *Written Communication,* 7 *(2),* 232-255.

Feez, S. & Joyce, H. (1998). *Text-based syllabus design.* Sydney: National Centre for English Language Teaching and Research.

Feez, S. (2001). Heritage and innovation in second language education. In A. Johns (Ed.), *Genre in the classroom: Multiple perspectives* (pp.43-69). Mahwah, NJ: Lawrence Erlbaum.

Ferris, D. (1994). Rhetorical strategies in student persuasive writing: Differences between native and non-native English Speakers. *Research in the Teaching of English, 28 (1),* 45-65.

Flower, L. (1979). Writer-based prose: A Cognitive basis for problems in writing. *College English, 41,* 19-37.

Flowerdew, J. (1993). An Educational, or process, approach to the teaching of professional genres. *ELT Journal, 47 (4),* 305-16.

Flowerdew, L. (1999). Critical thinking development and academic writing for engineering students. In M. Pally (Ed.), *Sustained content teaching: Developing academic skills in ESL/EFL writing classes.* Boston, MA: Houghton Mifflin.

Flowerdew, L. (2000). Using a genre-based framework to teach organizational structure in academic writing. *ELT Journal, 54 (4),* 369-78.

Freedman, A. & Medway, P. (1994). Introduction: New views of genre and their implications for education. In A. Freedman & P. Medway (Eds.), *Learning and teaching genre.* Portmouth, NH: Boynton/Cook Publishers-Heinemann.

Freedman, A. & Medway, P. (Eds.). (1994). *Learning and teaching genre.* Portsmouth, NH: Boynton/Cook Publishers-Heinemann.

Freedman, A. (1993a). Show or tell? The Role of explicit teaching in the learning of new genres. *Research in the Teaching of English, 27,* 222-51.

Freedman, A. (1993b). Situating genre: A Rejoinder. *Research in the Teaching of English, 27,* 272-8.

Fulcher, G. (1998). Assessing Writing. In G. Fulcher (Ed.), *Writing in the English language classroom.* Hertfordshire, UK: Prentice-Hall Europe ELT.

Fulkerson, R. (1996). *Teaching the argument in writing.* Urbana, IL: National Council of Teachers of English.

Grabe, W. & Kaplan, R. (1996). *Theory & practice of writing.* Harlow: Longman.

Graves, D. (1983). *Writing: Teachers and children at work.* Exeter: Heinemann.

Gray, B. (1987). How is natural language teaching—Employing wholistic methodology in the classroom. *The Australian Journal of Early Childhood, 12,* 4.

Green, J. (1992). *Making the links.* Melbourne: Adult Migrant Education Service. Video and Workshop Guide.

Green, S., Salkind, N. & Akey, T. (1999). *Using SPSS® for Windows®: Analyzing and understanding data.* Upper Saddle River, NJ: Prentice Hall.

Hammond, J., Burns, A., Joyce, H., Brosnan, D., & Gerot, L. (1992). *English for social purposes: A Handbook for teachers of adult literacy.* Sydney: National Centre for English Language Teaching and Research.

Hamp-Lyons, (1990). Second language writing: Assessment Issues. In B. Kroll (Ed.), *Second language writing: Research insights for the classroom* (pp. 69-87). Cambridge: Cambridge University Press.

Hamp-Lyons, L (1991a). Basic concepts. In L. Hamp-Lyons (Ed.), *Assessing second language writing in academic contexts.* Norwood, NJ: Ablex.

Hamp-Lyons, L. (1991). *Assessing second language writing in academic contexts.* Norwood, NJ: Ablex.

Harris, D. (1968). *Testing English as a second language.* New York: McGraw Hill.

Hasan, (1989). The Structure of text. In M. Halliday & R. Hasan, *Language,context, and text* (pp. 52-69). Oxford: Oxford University Press.

Hatch, E. & Lazaraton, A. (1991). *The Research manual: Design and statistics for applied linguistics.* Boston, MA: Heinle & Heinle.

Hawkins, B. (1988). *Scaffolded classroom interaction and its relation to second language acquisition for language minority children.* Ph.D. dissertation, University of California, Los Angeles.

Henry, A. & Roseberry, R. (1998). An Evaluation of a genre-based approach to the teaching of EAP/ESP writing. *TESOL Quarterly, 32* (1), 147-156.

Hergenhahn, B. R. & Olson, M. H. (1993). *An Introduction to theory of learning.* Englewood Cliffs, NJ: Prentice Hall.

Hillocks, G. (1986). *Research on written composition: New directions for teaching.* Urbana, IL: National Council of the Teachers of English.

Hinds, J. (1987). Reader versus writer responsibility: A New typology. In U. Connor & R. Kaplan (Eds.), *Writing across languages: Analysis of L2 text* (pp. 141-152). Reading, MA: Addison-Wesley.

Hinds, J. (1990). Inductive, deductive, quasi-inductive: Expository writing in Japanese, Korean, Chinese, and Thai. In U. Connor & A. Johns (Eds.), *Coherence in*

Hopkins, K. (1998). *Educational and psychological measurement and evaluation.* Needham Heights, MA: Allyn & Bacon.

Hughes, A. (1989). *Testing for language teachers.* Cambridge: Cambridge University Press.

Huot, B. (1996). Toward a theory of writing assessment. *College Composition and Communication, 47,* 549-66.

Hyland, K. (1990). A Genre description of the argumentative essay. *RELC Journal, 21,* 66-78.

Hyland, K. (2000). Disciplinary discourse: Social interactions in academic writing. Harlow: Longman.

Hyland, K. (2002). *Teaching and researching writing.* Harlow: Longman

Hyland, K. (2003a). Genre-based pedagogies: A Social response to process. *Journal of Second Language Writing, 12,* 17-29.

Hyland, K. (2003b). *Second language writing.* Cambridge: Cambridge University Press.

Hyon, S. (1995). *A Genre-based approach to ESL reading: implications for North America and Australia.* Unpublished doctoral dissertation, University of Michigan, Ann Arbor.

Hyon, S. (2001a). Genre and ESL reading: a classroom study. In A. Johns (Ed.), *Genre and pedagogy: Multiple perspectives.* Mahwah, NJ: Lawrence Erlbaum.

Hyon, S. (2001b). Long-term effects of genre-based instruction: A follow-up study of an EAP reading course. *English for Specific Purposes, 20*, 417-438.

Intaraprawat, P. & Steffenson, M. S. (1995). The Use of metadiscourse in good and poor ESL essays. *Journal of Second Language Writing, 4,* 253-272.

Intaraprawat, P. (2002). *Writing an argumentative essay.* Nakorn-radchasima: Suranaree University of Technology Press.

Jacobsen, D., Eggen, P., & Kanchak, D. (1999). *Methods for teaching: Promoting student learning.* Upper Saddle Rivers, NJ: Prentice Hall.

Jacoby, S., Leech, D., & Holten, C. (1995). A Genre-based developmental writing course for undergraduate ESL science majors. In D. Belcher & G. Braine (Eds.), *Academic writing in a second language.* Norwood, NJ: Albex Publishing.

Johns, A. (1997). *Text, role, and context: Developing academic literacies.* Cambridge: Cambridge University Press.

Johns, A. (1999). Opening our doors: Applying socioliterate approaches (SA) to language minority classrooms. In L. Harklau, K. Losey, & M. Siegal (Eds.), *Generation 1.5 meets college composition: Issues in the teaching of writing to U.S.-educated learners of ESL* (pp. 159-171). Mahwah, NJ: Lawrence Erlbaum.

Kaplan, R. (1966). Cultural thought patterns in intercultural education. *Language Learning, 16,* 1-20.

Kay, H. & Dudley-Evans, T. (1998). Genre: what teachers think. *ELT Journal, 52*, 308-314.

Kinneavy, J. (1971). *A Theory of discourse.* New York: Norton.

Kirkpatrick, L. & Feeney, B. (1998). *A Simple guide to SPSS® for Windows®.* Pacific Grove, CA: Brooks/Cole Publishing.

Knudson, R. (1992a). Analysis of argumentative writing at two grade levels. *Journal of Educational Research, 85 (3),* 169-179.

Knudson, R. (1992b). The Development of written argumentation: An Analysis and comparison of argumentative writing at four grade levels. *Child Study Journal, 22 (3),* 167-184.

Kowal, M. & Swain, M. (1994). Using collaborative language production tasks to promote students' language awareness. *Language Awareness, 3,* 73-93.

Krashen, S. (1984). *Writing: Research, theory and applications.* Torrance, CA: Laredo Publishing.

Kroll, B. (1990a). The Rhetoric-syntax split: Designing a curriculum for ESL students. *Journal of Basic Writing, 9,* 40-55.

Kroll, B. (1990b). What does time buy? ESL student performance on home versus class compositions. In B. Kroll (Ed.), *Second language writing: Research insights for the classroom* (pp. 140-154). Cambridge: Cambridge University Press.

Kroll, B. (1990c). Understanding TOEFL's Test of Written English. *RELC Journal, 22,* 20-33.

Kroll, B. (Ed.). (1990). *Second language writing: Research insights for the classroom.* Cambridge: Cambridge University Press.

Kumaravadivelu, (2001). Toward a postmethod pedagogy. *TESOL Quarterly,* 35 *(4),* 537-560.

Langan, J. (1996). *College writing skills.* NY: McGraw-Hill.

Larsen-Freeman, D. & Long, M. H. (1991). *An Introduction of second language acquisition research.* London: Longman.

Larson, C. (1998). *Persuasion: Reception and responsibility.* Belmont, CA: Wadsworth.

LeCompte, M. D. & Preissle, J. (1993). *Ethnography and qualitative design in educational research.* San Diego, CA: Academic Press.

Leki, I. & Carson, J. (1997). Completely different worlds: EAP and the writing experiences of ESL students in university courses. *TESOL Quarterly,* 31 *(1),* 39-70.

Lennon, P. (1991). Error: Some problems of definition, identification, and distinction. *Applied Linguistics,* 12 *(2),* 180-196.

Likhasith, A. (1995). *Effects of reading teaching methods using the genre-based approach for developing story writing skill and interest of Matayom Suksa I students at Ban Samrong Kokpet school in Surin province.* Unpublished M.Ed. thesis, Srinakharinwirot University, Bangkok.

MacArthur, C. & Ferretti, R. (1997, March). *The Effects of elaborated goals on the argumentative writing of students with learning disabilities and their normally achieving peers.* Poster session presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Macken, M. & Slade, D. (1993). Assessment: A Foundation for effective learning in school context. In B. Cope & M. Kalantzis (Eds.), *The Power of literacy: A Genre approach to teaching writing* (pp. 203-230). Pittsburgh: University of Pittsburgh Press.

Macken-Horarik, M. (2001). Something to shoot for: A Systemic functional approach to teaching genre in secondary school science. In A. Johns (Ed.), *Genre in the classroom: Multiple perspectives* (pp. 17-42). Mahwah, NJ: Lawrence Erlbaum.

Mason, J. E. & Bramble, J. W. (1997). *Research in education and the behavioral sciences: Concepts and methods.* Madison, WI: Brown & Benchmark.

McCann, T. (1989). Student argumentative writing knowledge and ability at three grade levels. *Research in the Teaching of English,* 23 *(1),* 62-76.

McNamara, T. (1996). *Measuring second language performance.* London: Longman.

Meredith, D. G., Borg, W. R., & Gall, J. P. (1996). *Educational research: An Introduction.* White Plains, NY: Longman.

Mitchell, S. (1994). *The Teaching and learning of argument in sixth forms and higher education: Final report.* Hull: University of Hull, Centre for Studies in Rhetoric.

Morrison, J. E. (Ed.). (1991). *Training for performance: Principles of applied human learning.* Chichester, West Sussex: John Wiley & Sons.

Murray, D. (1982). *Learning by teaching: Selected articles on writing and teaching.* Montclair, NJ: Boynton/Cook.

Mustafa, Z. (1995). The Effect of genre awareness on linguistic transfer. *English for Specific Purposes, 14*, 247-256.

Nadell, J., Langan, J., & McMeniman, L. (1993). *The Macmillan reader.* NY: Macmillan.

Nattinger, J. R., & DeCarrico, J. S. (1992). *Lexical phrases and language teaching.* Oxford: Oxford University Press.

Nicholas, J., & Nicholl, J. (1994). *Models for effective writing.* Needham Heights, MA: Allyn & Bacon.

Nipitkul, S. (1995). *An Investigation into the effects of translation and direct composition on students writing proficiency and their attitudes.* Unpublished master thesis, Naresuan University, Phitsanulok, Thailand.

North, B. & Schneider, G. (1998). Scaling descriptors for language proficiency scales. *Language Testing, 15* (2), 217-263.

Nunan, D. (1992). *Research methods in language learning.* Cambridge: Cambridge University Press.

Nunnally, J. (1978). *Psychometric methods.* New York: McGraw Hill.

O'Brien, T. (1995). Rhetorical structure analysis and the case of the inaccurate, incoherent source hopper. *Applied Linguistics,* 16 *(4),* 442-82.

Ohta, A. M. (2000). Rethinking interaction in SLA: Developmentally appropriate assistance in the zone of proximal development and the acquisition of L2 grammar. In J. P. Lantolf (Ed.), *Sociocultural Theory and Second Language Learning.* Oxford: Oxford University Press.

Ormrod, E. J. (2000). *Educational psychology.* Upper Saddle Rivers, NJ: Merrill Prentice Hall.

Ostler, S. (1987). English in parallels: A Comparison of English and Arabic prose. In U. Connor & R. Kaplan (Eds.), *Writing across language: Analysis of L2 text (pp. 169-185).* Reading, MA: Addison-Wesley.

Paltridge, B. (2001). *Genre and the language learning classroom.* Ann Arbor: University of Michigan Press.

Phuwichit, P. (2003). *An Analysis of the argumentative pattern written by Thai English major college students.* Unpublished master thesis, Naresuan University, Phitsanulok, Thailand.

Pollitt, A. & Murray, N. (1996). What raters really pay attention to. In M. Milanovic & N. Saville (Eds.), *Studies in language testing 3: Performance testing, cognition, and assessment* (pp. 74-91). Cambridge: Cambridge University Press.

Pollitt, A. (1991). Giving students a sporting chance: Assessment by counting and by judging. In C. Alderson & B. North (Eds.), *Language testing in the 1990s* (pp. 46-59). Hemel Hempstead, Hertfordshire: Phoenix ELT/Prentice-Hall UK.

Prabhu, N. (1990). There is no best method—why? *TESOL Quarterly, 24,* 161-176.

Pringle, I. & Freedman, A. (1985). *A Comparative study of writing abilities in two modes at the grade 5, 8, and 12 levels.* Toronto: Ministry of Education.

Raimes, A. (1991). Out of the woods: Emerging traditions in the teaching of writing. *TESOL Quarterly, 25 (3),* 407-30.

Raphael, T., Englert, C., & Kirschner, B. (1989). Students' metacognitive Knowledge about Writing. *Research in the Teaching of English, 23 (4),* 343-79.

Reid, J. (1988). *The Process of composition.* Englewood Cliffs, NJ: Prentice Hall Regents.

Reid, J. (1993). *Teaching ESL writing.* Englewood Cliffs, NJ: Prentice Hall Regents.

Reid, S. (1989). *The Prentice Hall guide for college writers.* Englewood Cliffs, NJ: Prentice Hall.

Reid, S. (1991). *Purpose and process: A Reader for writers.* Englewood Cliffs, NJ: Prentice Hall.

Reppen, R. (1995). A Genre-based approach to content writing instruction. *TESOL Journal, 4* (2), 32-35.

Reznitskaya, A. (2002). *The Influence of group discussions and explicit instruction on the acquisition and transfer of argumentative knowledge.* Unpublished doctoral dissertation, University of Illinois at Urbana-Champaign, Illinois.

Richard-Amato, P. (2003). *Making it happen: From interactive to participatory language teaching: Theory and practice.* (3rd ed.). White Plains, NY: Longman.

Rogoff, B. (1990). *Apprenticeship in thinking: Cognitive development in social context.* New York: Oxford University Press.

Rosenshine, B. & Guenther, J. (1992). Using scaffolds for teaching higher level cognitive strategies. In J. W. Keefe & H. J. Walberg (Eds.), *Teaching for Thinking.* Reston, VA: National Association of Secondary School Principles.

Sawyer, W. & Watson, K. (1987). Questions of genre. In I. Reid (Ed.), *The Place of genre in learning: Current debates.* Geelong, Victoria: Deakin University Press.

Seedokmai, S. (1999). *An Analysis of the rhetorical pattern applied by Srinakharinwirot university second-year English majors.* Unpublished master thesis, Bangkok, Thailand.

Seliger, H. & Shohamy, E. (1989). *Second language research methods.* Oxford: Oxford University Press.

Shannon, D. & Davenport, M. (2001). *Using SPSS® to solve statistical problems: A Self-instruction guide.* Upper Saddle River, NJ: Merrill Prentice Hall.

Shaughnessy, M. (1977). *Errors & expectations.* New York: Oxford University Press.

Silva, T. (1993). Toward an understanding of the distinct nature of L2 writing: The ESL research and its implications. *TESOL Quarterly, 27 (4),* 657-77.

Smalley, R., Ruetten, M, &Kozyrev, J. (2001). *Refining composition skills: Rhetoric and grammar.* Boston, MA: Heinle & Heinle.

Snowman, B. (1993). *Psychology applied to teaching.* Boston, MA: Houghton Mifflin.

Stein, L. & Glenn, C. (1979). An Analysis of story comprehension in elementary school children. In R. Freedle (Ed.), *New directions in discourse processing,* Vol. 2 in the series *Advances in Discourse Processes.* Norwood, NJ: Ablex.

Swales, J. & Feak, C. (1994). *Academic writing for graduate students: Essential tasks and skills.* Ann Arbor: University of Michigan Press.

Swales, J. & Feak, C. (2000). *English in today's research world: A Writing guide.* Ann Arbor: University of Michigan Press.

Swales, J. (1990a). *Genre analysis: English in academic and research settings.* Cambridge: Cambridge University Press.

Swales, J. (1990b). Nonnative speaker graduate engineering students and their introductions: Global coherence and local management. In U. Connor & A. Johns (Eds.), *Coherence in writing: Research and pedagogical perspectives.* Washington, DC: TESOL.

Swales, J. (2000). *Further reflections on genre and ESL academic writing.* Abstract of the keynote presentation to the Symposium on Second Language Writing, Purdue University, Lafayette, IN, September.

Thai theses. (2003). [CD-ROM]. TIAC.

Tharp, R. G. & Gallimore, R. (1991). *The Instructional conversation: Teaching and learning in social activity.* New York: Cambridge University Press.

Thornburg, D. (1991). Strategy instruction for academically at-risk students: An Exploratory study of teaching higher-order reading and writing in the social studies. *Reading, Writing, and Learning Disabilities, 7,* 377-406.

TIAC. (2003). *Thai theses online.* Available http://www.tiac.or.th.

Toulmin, S. (1958). *The Uses of argument.* Cambridge: Cambridge University Press.

Toulmin, S., Rieke, R., & Janik, A. (1979). *An Introduction to reasoning.* New York: Macmillan.

Tribble, C. (1996). *Writing.* Oxford: Oxford University Press.

Van Eemeren, F., Grootendorst, R., Henkemans, F., Blair, J., Johnson, R., Krabbe, E., Plantin, C., Walton, D., Willard, C., Woods, J., & Zarefsky, D. (1996). *Fundamentals of argumentation theory: A Handbook of historical backgrounds and contemporary developments.* Mahwah, NJ: Lawrence Erlbaum.

Vande Kopple, W. J. (1985). Some exploratory discourse on metadiscourse. *College Composition and Communication, 36,* 82-93.

Varghese, S. & Abraham, S. (1998). Undergraduates arguing a case. *Journal of Second Language Writing, 7* (3), 287-306.

Vesterman, W. (2000). *Reading and writing short arguments.* Mountain View, CA: Mayfield.

Vygotsky, L. (1978). *Mind in society: The Development of higher psychological processes.* Cambridge: Cambridge University Press.

Walberg, H. J. (1990). Productive teaching and instruction: Assessing the knowledge base. *Phi Delta Kappan, 71,* 470-478.

Ward, J. (1999). How large a vocabulary do EAP engineering students need? *Reading in a Foreign Language, 12,* 309-324.

Ward, J. (2001). EST: evading scientific text. *English for Specific Purposes, 20,* 141-152.

Ward, J. (2004). Estimating SUT students' vocabulary size. *English Language Studies Forum, 1,* 2004.

Webb, S. (1994). *The Resourceful writer: Readings to accompany the Harbrace college handbook.* Forth Worth, TX: Harcourt Brace.

Weigle, S. (2002). *Assessing writing.* Cambridge: Cambridge University Press.

Weissberg R. & Buker, S. (1990). *Writing up research: Experimental report writing for students of English.* Englewood Cliffs, NJ: Prentice Hall Regents.

Wells, G. (1999). *Dialogic inquiry: Toward a sociocultural practice and theory of education.* Cambridge: Cambridge University Press.

White, E. (1984). Holisticism. *College Composition and Communication, 35 (4)*, 400-409.

White, E. (1994). *Teaching and assessing writing: Recent advances in understanding, evaluating and improving student performance.* San Francisco: Jossey-Bass.

White, J. (1989). Children's argumentative writing: A Reappraisal of difficulties. In F. Christie (Ed.), *Writing in schools: Reader* (ECT 418) (pp. 9-23). Geelong, Victoria: Deakin University Press.

Widdowson, H. (1990). *Aspects of language teaching.* Oxford: Oxford University Press.

Widdowson, H. G. (1983). *Learning purpose and language use.* Oxford: Oxford University Press.

Wilkinson, A., Barnsley, G., Hanna, P., & Swan, M. (1980). *Assessing language development.* Oxford: Oxford University Press.

Williams, J. M. & Colomb, G. G. (1993). The Case of explicit teaching: Why what you don't know won't help you. *Research in the Teaching of English, 27,* 252-64.

Wong, H. (1988). *The Use of rhetorical questions in written argumentative discourse.* Unpublished manuscript, Georgetown University, Washington, DC.

Wood, N. (1998). *Writing argumentative essays.* Upper Saddle River, NJ: Prentice Hall.

Yeh, S. (1998a). Validation of a scheme for assessing argumentative writing of middle school students. *Assessing Writing, 5* (1), 123-150.

Yeh, S. (1998b). Empowering education: Teaching argumentative writing to cultural minority middle-school students. *Research in the Teaching of English, 33*, 49-83.

Yoshimura, T. (2002). *Formal instruction of rhetorical patterns and the effectiveness of using the L1 in argumentative writing in an EFL Setting.* Unpublished doctoral dissertation, Temple University, Osaka, Japan.

APPENDIX A

Holistic Rubric for Scoring Student Argumentative Writing
(Adapted from Oregon Dept. of Education's Student Language Scoring Guide 2003-4)

Scale Score	Rhetorical Control	Language Control
5 Elaborated Argument	An essay at this level fulfils the writing task expectations successfully. A typical essay in this category: • clearly addresses the topic with thoughtful thesis • is well focused, organized, and developed with effective and appropriate use of transitions • demonstrates thorough understanding of the issues presented; extensive use of specific, well-developed data of a variety of types to support the thesis • acknowledges and responds to major objections adequately and effectively • provides effective and complete closure to the essay	• is clearly written with few errors; errors do not interfere with comprehension • includes academic vocabulary that is rarely inaccurate or repetitive • includes generally accurate word forms and verb tenses • uses a variety of sentence types accurately • contains source texts language that is well integrated with student-generated language
4 Developed Argument	An essay at this level fulfils the writing task expectations but not at the highest level as a 5 rated essay. A typical essay in this category: • addresses the topic with clear thesis • is generally well organized and developed, using effective and appropriate transitions • demonstrates competent understanding of the issue presented; extensive use of specific, well-developed data of a variety of types to support the thesis, but more detail may still be desirable • acknowledges and responds to major objections generally well • provides competent conclusion that reinforces and comments on the thesis	• is clearly written with few errors; errors do not interfere with comprehension • includes academic vocabulary that is rarely inaccurate or repetitive • may include inaccurate word forms and verb tenses • uses a variety of sentence types • incorporates ideas from readings or outside sources without plagiarism; most sources are documented correctly using varied styles
3 Fairly Developed Argument	An essay adequately meets the task expectations, though it may fulfil some parts of the task more effectively than the others. An essay in this category: • addresses the topic adequately with thesis, though it could have been more focused • is adequately organized and developed though development may be thin at times or some transitions among parts may be desirable • demonstrates adequate understanding of the topic presented; some variety in use of the data to support the thesis, though some supports are less compelling or could have been better developed • adequately acknowledges and responds to some important counter-arguments • provides conclusion that reinforces and comments on the thesis	• is generally clearly written with some errors; errors may interfere with comprehension occasionally • demonstrates occasional problems with word choice • includes some inaccurate word forms and verb tenses • uses a variety of sentence types with occasional errors • incorporates ideas from readings or outside sources; most sources are documented correctly

Scale Score	Rhetorical Control	Language Control
2 **Inadequately Developed Argument**	The task is attempted but not adequately achieved or only partially successful. An essay: • may not address the topic adequately or be sufficiently focused • may not be adequately organized or developed or is organized in parts but other parts are disjointed or lack transitions; may lack development in parts • may demonstrate lack of understanding of the issue presented; may be illogical or have insufficient or inappropriate support for the thesis; may use limited range of information to support the thesis; may fail to cite sources of ideas or quotations • takes into account somewhat important objections which may not be responded adequately • provides conclusion summarizing the main parts but may neither reinforce nor comment on the thesis	• contains many errors; some errors may interfere with comprehension • includes limited vocabulary or examples of inappropriate word choice • includes a number of inaccurate word forms • contains some problems with verb tenses • uses limited types of sentences • may not incorporate ideas from readings or outside sources without plagiarism; sources may not be cited correctly
1 **Attempted Argument**	The task is attempted but slightly fulfils the expectations. An essay: • does not deal adequately with the topic; may be unclear or poorly focused • may have serious problems with organization and development; some parts may be missing or underdeveloped; has few or no transitions among parts • demonstrates lack of understanding of the issue presented; may have irrelevant specifics or unsupported generalizations; supports lack in amount or variety or both; may fail to cite sources of ideas or quotations • takes into account less important objections with responses given but not seriously • provides conclusion that summarizes the main parts but neither reinforce nor comment on the thesis	• contains numerous errors • contains errors that often interfere with comprehension • uses simple and repetitive vocabulary that may not be appropriate for academic writing • uses inappropriate word forms and verb tenses • does not vary sentence types sufficiently • does not incorporate ideas from readings or outside sources without plagiarism; most sources are not cited correctly
0 **Off Task**	An essay at this level fails to fulfil the writing task expectations. A student writer has not produced an assessable argumentative essay. A paper is rated 0 if it reads as non-argumentative prose or reveals one or more of the following weaknesses: • is descriptive • deals with the issue which is not arguable or can only be viewed as facts • is severely underdeveloped or contains inadequate, incoherent, or illogical presentation of ideas that does not orient the reader sufficiently to the topic or thesis • contains severe and persistent errors that interfere with understanding throughout the reading • demonstrates serious disregard of English writing conventions including correct spelling, capitalization, punctuation, paragraphing, and documentation of information source • contains no in-text citations or references	

APPENDIX B

Analytic Rubric for Scoring Student Argumentative Writing
(Adapted from Oregon Dept. of Education's Student Language Scoring Guide 2003-4)

CLAIM	5	The introduction orients the reader sufficiently and effectively to the topic and to the author's thesis.
	4	The introduction addresses the topic clearly but may not be as effective as a 5-rated paper in this category.
	3	The introduction addresses the topic clearly. Its parts are developed and organized in a way that orients the reader adequately to the topic and to the author's thesis, though some parts of the paragraph could have been better developed or elaborated.
	2	The introduction may be brief or inadequately developed and not orient the reader sufficiently to the topic and to the author's thesis, rendering the proposed solution somewhat practical and the problem somewhat real. Topic should be more focused with respect to audience and purpose. Some elements in the paragraph are underdeveloped.
	1	The introduction is brief and less than adequate and does not orient the reader sufficiently to the topic and to the author's thesis, making the feasibility of the proposed solution highly questionable or raising only superficial argument. Problem may not exist at all. Topic lacks clarity and focus and some elements are missing or undeveloped.
	0	The introduction lacks clarity and focus and does not orient the reader to the topic and to the arguable thesis.
REASON	5	Supporting points and details are exceptionally well chosen and appropriate to audience and purpose and in amount and variety. Details are clearly and insightfully developed and organized providing the strongest possible justification for the claim. Where appropriate, use of resources provides strong, accurate, and credible support to the claim. Sources are well integrated and documented.
	4	Supporting points are strong, relevant, and important. Developmental details are well chosen, developed, and appropriate in amount and variety though may not be as effective as a 5-rated paper in this category. When needed, use of resources provides strong, accurate, and credible support to the claim. Sources are generally well integrated and documented.
	3	Support is relevant, important, and appropriate in amount and variety though some of which, relative to other reasons, may be less compelling or not strongly related to the claim. Some supporting details could have been better developed or made clearer. Documentation is used properly most of the time when appropriate.
	2	Support is relevant but some points may be weak, or unimportant. Developmental details may occasionally be descriptive, less than adequate, or not be varied enough. Some selected details are perhaps not consistently well chosen for audience and purpose, and may not be based on reliable sources. Documentation is sometimes used to cite sources of information.
	1	Support is attempted but developmental details are often limited in variety and amount, uneven, predictable, irrelevant, or unimportant. Most details may not be well grounded in credible resources; they may be based on biased or uninformed generalizations, or questionable sources of information. Documentation is frequently neglected or not used properly when appropriate.
	0	All the reasons stated are overly broad or simplistic, predictable, irrelevant or not grounded in credible resources.

(Rubric continues on next page)

REBUTTAL TO COUNTER-ARGUMENT	5	Important counter-arguments are acknowledged and responded adequately and effectively.
	4	Important counter-arguments are generally well acknowledged and responded.
	3	Some important counter-arguments are stated and refuted adequately.
	2	Somewhat important counter-arguments are stated but may not be refuted adequately.
	1	Some relatively unimportant counter-arguments are stated but may not be refuted adequately.
	0	No possible counter-argument is identified.

APPENDIX C

Student Writing Questionnaire

Notes: This questionnaire was administered to the experimental group with the returning rate of 65% (19 out of 29 students). Response frequencies were reported in percentage. Written responses to open prompt in Part III were returned from only 8 student respondents.

Part I: Writing Background (6 questions)

1. What kinds of English writing did you do before coming to SUT? Please indicate. (Note: Each student may indicate more than one type of writing)

   - diaries/journals/letters *(27%)*
   - summaries/reflections on assigned readings *(23%)*
   - writing freely about a certain topic/event *(88%)*
   - short answers to comprehension questions *(95%)*
   - short messages/stories *(46%)*

2. Please estimate the amount of required writing that you did in high school?

   2.1 At a paragraph level
       2.1.1 about one paragraph *(89%)*
       2.1.2 more than one paragraph *(11%)*

   2.2 At a full-length essay level *(None)*

3. What do you feel when writing in English?

   a) I like it. *(5%)*
   b) I quite like it. *(26%)*
   c) I write because I have to. *(32%)*
   d) I don't like it. *(37%)*
   e) I don't like it very much. *(None)*

4. How important is English for your future?

   a) very important *(55%)*
   b) quite important *(36%)*
   c) not quite important *(9%)*
   d) not important at all *(None)*

5. How difficult is it for you to write in English?

   a) very difficult *(10%)*
   b) quite difficult *(60%)*
   c) not quite difficult *(20%)*
   d) not difficult at all *(10%)*

6. How often do you have trouble with the followings while writing in English?

   6.1 using correct grammar
      a) very often *(42%)*
      b) often *(37%)*
      c) sometimes *(16%)*
      d) almost never *(5%)*

   6.2 finding right words to express ideas
      a) very often *(32%)*
      b) often *(42%)*
      c) sometimes *(26%)*
      d) almost never *(None)*

   6.3 getting good ideas to write on
      a) very often *(37%)*
      b) often *(32%)*
      c) sometimes *(31%)*
      d) almost never *(None%)*

   6.4 organizing my ideas in a clear way
      a) very often *(11%)*
      b) often *(53%)*
      c) sometimes *(31%)*
      d) almost never *(5%)*

Part II: Writing in English 5 (1 question)

1. How important did you perceive of the following aspects of writing while composing your different versions of argumentative essay?
(Please rank by indicating 1 = least important; 2 = little important; 3 = quite important; and 4 = very important)

   1.1 first draft

      1.1.1 using correct grammar
         (very important = 0%; quite important = 6%; little important = 37%; least important = 47%)

    1.1.2   using appropriate vocabulary
           (very important = 0%; quite important = 37%; little important = 32%; least important = 31%)

    1.1.3   developing ideas/contents
           (very important = 79%; quite important = 10%; little important =11%; least important = 0%)

    1.1.4   organizing ideas/contents
           (very important = 11%; quite important = 47%; little important = 21%; least important = 21%)

1.2 subsequent draft(s)

    1.2.1   using correct grammar
           (very important = 21%; quite important = 16%; little important = 32%; least important = 31%)

    1.2.2   using appropriate vocabulary
           (very important = 0%; quite important = 30%; little important = 30%; least important = 40%)

    1.2.3   developing ideas/contents
           (very important = 42%; quite important = 21%; little important = 16%; least important = 21%)

    1.2.4   organizing ideas/contents
           (very important = 32%; quite important = 37%; little important = 26%; least important = 5%)

1.3 final draft

    1.3.1   using correct grammar
           (very important = 47%; quite important = 26%; little important = 21%; least important = 6%)

    1.3.2   using appropriate vocabulary
           (very important = 0%; quite important = 32%; little important = 21%; least important = 47%)

    1.3.3   developing ideas/contents
           (very important = 26%; quite important = 16%; little important = 37%; least important = 21%)

    1.3.4   organizing ideas/contents
           (very important = 26%; quite important = 26%; little important = 21%; least important = 7%)

Part III: Students' comments (1 open prompt)

You are encouraged to freely express your opinions about the course in terms of what you like and dislike, how helpful the lessons are, suggestions you have, etc. (See 4.4 for summarized students' written comments)

Vita

Wiwat Udomyamokkul was born in Bangkok on September 2, 1972. In 1995, he graduated from Assumption University (formerly called Assumption Business Administration College [ABAC]) with a Bachelor of Business Administration in International Business Management minoring in Accounting. He was awarded a Postgraduate Diploma in International Business Management from London South Bank University in 1999. After graduation in 1995, he worked for a multinational multimodal transporter firm in Bangkok, as a sales executive in the import-export department responsible for customer relations, customs procedures, selling, and after-sales activities.

After moving to Korat in 2001, he joined the Master of Arts program in English language studies offered by Suranaree University of Technology (SUT). While studying at SUT, he was given an opportunity to serve as a teaching assistant as a part of his study and research. His special interests in TEFL are in the fields of teaching writing, writing assessment, and genre-based methodology.