THE EFFECT OF TRAINING STUDENTS ON SELF-ASSESSMENT OF THEIR

WRITING

/0/996

A THESIS PRESENTED BY

İKLİL KAYA YILDIRIM

TO THE INSTITUTE OF ECONOMICS AND SOCIAL SCIENCES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF MASTER OF ARTS

IN TEACHING ENGLISH AS A FOREIGN LANGUAGE

101996

BILKENT UNIVERSITY

JULY, 2001

ABSTRACT

Title: The Effect of Training Students on Self-Assessment of Their Own Writing

Author: İklil Kaya Yıldırım

Thesis Chairperson: Dr. James C. Stalker
                      Bilkent University, MA TEFL Program

Committee Members: Dr. Hossein Nassaji
                   Dr. William E. Snyder
                   Bilkent University, MA TEFL Program

This study aimed to investigate the effect of training students to self-assess their own writing. The study basically sought to find out (1) the effect of training students to self-assess their own writing, on the quality of self-assessment, and (2) the effect of training students to self-assess their own writing, on their writing skills. The study was conducted in the First Year English Program (FYE) at Bilkent University where the students receive content-based instruction courses and practice process writing.

The participants of the study were 25 Bilkent University Freshman Engineering and Science students. There were two groups: one treatment and one control. In the treatment group there were 13 students and in the control group there were 12 students. The students in the treatment group were given training on how to self-assess their own writing. The students in the control group self-assessed their writing, without any training, during the course of their usual instruction. After the training, the students in the treatment group were also administered an attitude questionnaire to elicit their thoughts about the effectiveness of training and about the practice of self-assessment.

The analysis of the data indicated that the students in the treatment group seemed to have improved their self-assessment skills consistently throughout the

training and got closer to the scorings of the external raters while the students in the control group were observed to be less consistent in their ratings. Also, although there was significant improvement in the writing skills of the students in both groups throughout the writing process, no statistically significant difference was observed between the two groups' in terms of writing improvement. Finally, the analysis of the data collected through the attitude questionnaire showed that most of the students in the treatment group perceived the training as effective and their attitudes toward self-assessment were in general positive.

The results indicate that the training had been effective particularly on the quality of self-assessment. Thus, it would appear that the instructors in the First Year English Program, at Bilkent University could benefit from the findings of the current study as it yielded encouraging results for engaging the students in their own learning and assessment process.

BILKENT UNIVERSITY

INSTITUTE OF ECONOMICS AND SOCIAL SCIENCES

MA THESIS EXAMINATION RESULT FORM

July 31, 2001

The examining committee by the Institute of Economics and Social Sciences for the

thesis examination of the MA TEFL student

İklil Kaya Yıldırım

has read the thesis of the student.

The committee has decided that the thesis of the student is satisfactory.
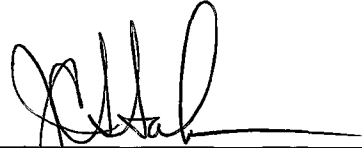
Thesis Title: The Effect of Training Students on Self-Assessment of Their Writing

Thesis Advisor: Dr. Hossein Nassaji
                        Bilkent University, MA TEFL Program

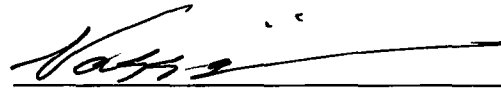Committee Members: Dr. James C. Stalker
                        Bilkent University, MA TEFL Program

                        Dr. William E. Snyder
                        Bilkent University, MA TEFL Program

We certify that we have read this thesis and that in our combined opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Arts.

Dr. James C. Stalker
(Chair)

Dr. Hossein Nassaji
(Committee Member)

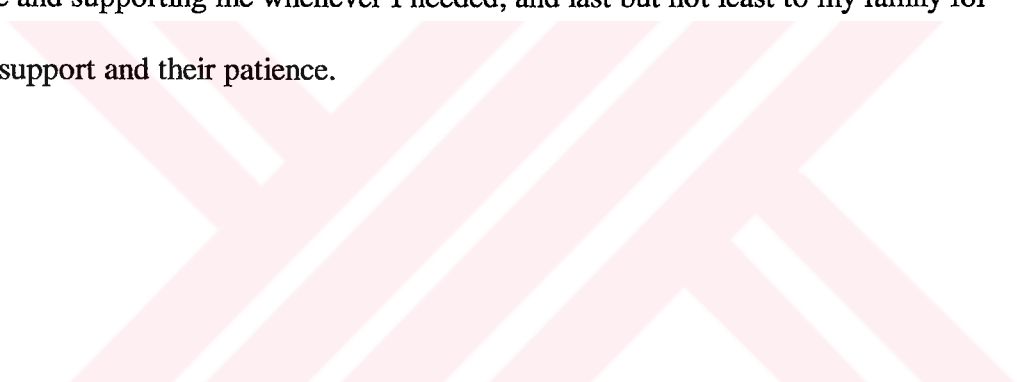Dr. William E. Snyder
(Committee Member)

Approved for the
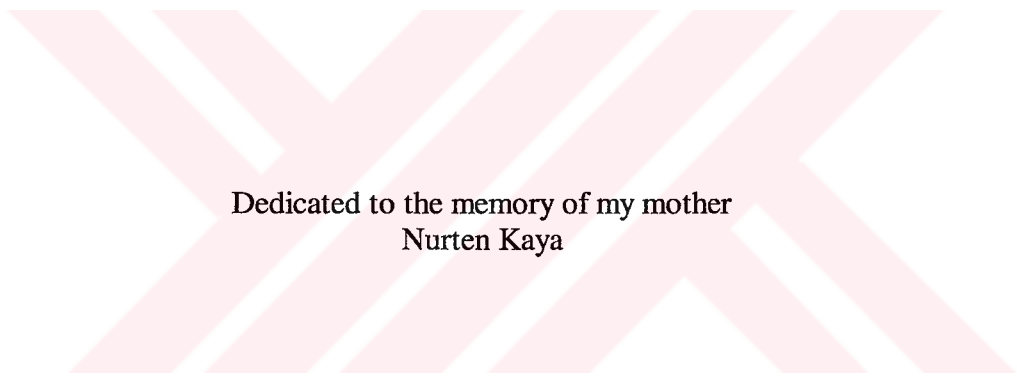Institute of Economics and Social Sciences

Kürşat Aydoğan
Director
Institute of Economics and Social Sciences

ACKNOWLEDGEMENTS

Dedicated to the memory of my mother
Nurten Kaya

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 1: INTRODUCTION

Background of the Study

The aim of this research is to investigate the effect of training on students'

self-assessment of their writing, in the First Year English Program at Bilkent

University. The issue of assessment is, no doubt, one of the most important

components of language curriculum since it functions to determine the current

knowledge the learner possesses. After 1980s and 1990s along with the innovations

in language teaching, different methods of assessment have been introduced to the

field of language teaching which are useful for different purposes (Brown & Hudson,

1998). This search for alternative methods of assessment is a consequence of the

dissatisfaction with traditional assessments in monitoring and measuring the student

performance in new methods of language teaching.

One of the fields where alternative assessments have been used most often is

writing. Back in 1930s and 1940s, the shared method of assessment in writing was

direct assessment. Later, in 1950s and 1960s teachers and learners spent much of

their effort in writing classes getting prepared for multiple choice tests that served

mostly for college entry exams. Together with the increasing importance in teaching

language communicatively in 1970s, new methods of teaching, like task-based

learning started to be used in the 1980s. In the light of all these changes, the search

for more meaningful, reliable and valid ways of assessment started (Hamp-Lyons,

1993) and a lot of alternative assessments, including self-assessment were introduced

in the field with the purpose of helping the new educational objectives (Brown &

Hudson, 1998).

As defined by Brown, " Self-assessments are any assessments that require students to judge their own language abilities or language performance"(1998, p. 53). Brown highlights the advantages of self-assessment as having the quality to be directly integrated into the language learning/teaching process, allowing for individualized assessments for each learner, providing an ongoing assessment process parallel to that of learning process, not requiring extra time and resources, involving learners in the assessment process, encouraging learner reflection on the learning process and also learner autonomy, and creating a positive attitude in the learners toward their learning process.

Ekbatani and Pierson (2000) similarly suggest that self-assessment can be used as a tool to assess the learners' language abilities, the strategies they use in developing these abilities, and the efficiency of the learners' engagement in both the learning and assessment process. Beyond the advantages it offers, other recent changes in the structure of the language curriculum toward learner-centered language teaching that focus mainly on the involvement of learners in their own learning process is another factor stimulating the use of self-assessment, particularly in writing (Nunan, 1988).

However, despite all these positive qualities that alternative assessments have, several questions have been raised in the field about their reliability and validity (Brown & Hudson, 1998) and also their objectivity (Huerta-Macias, 1995). Brown (1998) refers to some important disadvantages of self-assessments, such as the relative subjectivity of scoring, possible variations in scoring in relation to the skill levels and the unreliability of the scores under high-stakes circumstances like final exams.

At the same time, however, several ways to improve the reliability and validity of these assessments have been suggested which include " credibility, auditability, multiple tasks, rater training, clear criteria, and triangulation of any decision-making procedures" (Brown and Hudson, 1998, p. 655). Huerta-Macias (1995) defines credibility of an alternative assessment instrument as truth value of the testing instrument, that is to say, whether it measures what it intends to measure, and defines auditability as the consistency of the results of measurement replicated under the same conditions. Wilde, Del Vecchio and Gustke (as cited in Huerta-Macias, 1995) suggest several other ways of ensuring reliability in alternative assessments including designing various tasks that will yield the same results, using trained readers, working with clear criteria and anchor papers, and monitoring to see whether the raters are using the criteria in consistent manner.

However, among all these, training the learner has also been suggested, and has received special attention, as a means of improving the validity of self-assessment (Brown & Hudson,1998). Dickinson (1993) explains why learners need to be trained on self-assessment and points out that training develops students' ability to monitor their own progress, to identify the problems in their paper and solve those problems, and to control their own writing process. O'Malley and Pierce (1996) also suggest training the learners on the grading criteria as an important means of improving self-assessment gradually.

Different methods have been suggested in the literature about how to train learners to self-assess themselves. Methods that are widely employed in the field are: having the students study good models of writing and having the students apply sets of criteria to their own writing or the writing of others (Hillocks, 1986).

It is suggested that "most of the studies involved found statistically significant differences between students using the sets of criteria and those taught through some other technique" (Hillocks, 1986, p. 156). Thus, one of the ways by which self-assessment may be made more reliable is by training the self-assessors and helping them to use certain criteria.

Considering the points discussed above, it is essential to think about ways to not only integrate learners' self-assessment in the EFL writing classes, but also to improve their validity. Such efforts may help the language instructors to raise awareness in learners about their performance in writing and demonstrate how this attempt may contribute to the improvement of their performance of the particular skill. Thus, this research was conducted to investigate the effect of training on students' self-assessment of their own writing, in the First Year English Program, at Bilkent University. The findings of this study may also contribute to the instructional and assessment practices of the language instructors in this institute.

Statement of the Problem

In the First Year English Department (FYE) at Bilkent University, the method of instruction is content based (CBI), in which students are required to produce essays based on the ideas presented in the reading texts they study during the course. The writing approach used by the department is process writing. Students write essays of different types in several drafts and receive constant feedback from their instructors throughout the writing process. The instructors in the First Year English Program at Bilkent University always encourage the students to use writing criteria while performing their writing assignments, to help them develop the ability to monitor their own writing performance. This is done to involve the students' in

their own learning process. The set of criteria used by the instructors is either placed in the students' course materials or handed in to the students together with assignment prompts.

One important problem the teachers have is students' insistent disregard of the criteria despite their instructors' continuously reminding them about how it may contribute to their writing. Moreover, even if students use the writing criteria as suggested by their instructors, they seem unable to apply it efficiently to their writing. They most often can not see the reason why the teacher assessed their final grade as lower than what they expected, when they think that they have covered all the aspects of the writing criteria. Even when the instructors explain why and how the students receive a particular grade with reference to the writing criteria, the students seem dissatisfied with the explanations. This situation raises the necessity of considering possible reasons for the problem like whether there are differences between the students' and the teachers' understanding of the criteria or whether the components of the writing criteria are clear enough to help the students understand its applications. With regard to these concerns, it seems possible that if the students are involved in a self-assessment process of their own writing and if they are trained to use the criteria, they not only will have a chance to develop a common understanding with the instructor about what has been intended with each component of the writing criteria, but can also improve their writing skills. Students' involvement in training on self-assessment of their own writing skills may also enable them to apply the criteria efficiently, to perform better and evaluate their own writing. Thus it may prevent the demotivation students are likely to develop towards writing classes.

Purpose of the Study

The purpose of this research is to explore the effects of training the First Year

English Program students at Bilkent University to self-assess their own writing using

the pre-set course criteria. The study also aims at finding out whether training the

students' to apply the criteria for self-assessment can improve their writing ability.

Research Questions

Does training students' to self-assess their writing, when they use the course

criteria, affect the quality of their self-assessment?

Does training students to self-assess their writing, when they use the course

criteria, improve their writing skills?

Significance of the Problem

This study will be the first attempt to explore the effects of training students

to self-assess their own writing in the First Year English Program, at Bilkent

University. So the study is unique in its own academic context. Moreover, according

to the FYE Principles set for year 2000, the teachers are expected to engage students

in their own reflective learning process through allowing them to actively contribute

to the process. In this context, two of the suggestions offered are to engage students

in the analysis and evaluation of a given assignment criteria and also to engage the

students in self-assessment as part of the course. In this sense, the results of this

research may yield useful information for the realization of these particular goals.

Moreover, training the students to evaluate their writing skills based on the

course criteria may eliminate students' frustration and disappointment caused by the

teacher-assessed final grade. By doing so, students may become more autonomous

and become aware of the strengths and weaknesses of their own writing.

CHAPTER 2: REVIEW OF LITERATURE

Introduction

This research is designed to explore the effects of training learners to self-assess their own writing. In this chapter, the related literature will be reviewed with particular emphasis on the effect of training on self-assessment. The first section reviews the literature on language assessment in general and writing assessment in particular with emphasis on the alternative assessment practices used in process writing approach. The second section discusses self-assessment, specifically the self-assessment of the writing skills and related research. The last section provides a review of research done in the field so far about training on self-assessment of writing skills.

Language Assessment

Assessment is an inherent part of the teaching process and its significance can best be understood by looking at its undeniably important functions and the purposes it serves in the field of ELT. Brown (1995) categorises types of decisions made based on assessment results as " proficiency, placement, achievement and diagnosis " (p. 137). Primarily, assessment provides valuable information about the achievement of educational goals and therefore helps critical consideration of instructional and curricular needs. Secondly, it helps the decision making process and formulation of educational policies. Thirdly, and maybe the most important of all, it helps to monitor students' progress and their level of performance, that is, the outcome of their learning process. With regard to the crucial role of assessment in the educational process, a lot of work has been undertaken and different approaches have been suggested for the achievement of the best possible educational results. A review

of the literature on language assessment reveals the enormous efforts made for these purposes.

In 1950s and 1960s multiple choice and true-false tests were the most commonly used types of assessment. However, in 1970s and early 1980s cloze tests and dictation were widely employed for assessment purposes. It was after late 1980s and in 1990s that the communicative approach in the field of ELT became highly influential and urged the need to search for more valid, reliable and meaningful approaches to assessment (Brown & Hudson, 1998). Together with the wide interest in the implementation of the learner-centered curriculum as the principle aim of a great number of ESL /EFL training institutions and classrooms (Nunan, 1988), many professionals started to explore meaningful ways and means to actively engage the learners in the language assessment process (Ekbatani & Pierson, 2000). The need for improved assessment procedures within the scope of applied linguistics then led to a strong demand for the use of innovative approaches to assessment. Hence alternatives to traditional approaches in assessment were introduced (Brown & Hudson, 1998).

Alternative assessment has been defined in many different ways. Stiggins (1991) defines alternative assessments as any method employed for determining the knowledge the learner has or can apply, and that is different from the traditional forms of assessment. Some examples to these methods are "portfolios, conferences, diaries, peer-assessment and self-assessment "(Brown & Hudson,1998, p. 657). Portfolios are compilations of any aspects of students' work demonstrating their achievements, skills, efforts and contributions to a particular course. Conferences involve the student meeting with the teacher to discuss a particular piece of work or

learning process. Peer assessment is when students are engaged in the evaluation of each others' works (O'Malley & Pierce,1998). Finally, self-assessment requires students to rate their own language performance. (Brown & Hudson, 1998).

Moreover, O'Malley and Pierce (1998) assert that "alternative assessment is by definition criterion-referenced and is typically authentic because it is based on activities that represent classroom and real-life settings"(p. 2). Bachman (1990) describes criterion-referenced tests as those " designed to enable the test user to interpret a test score with reference to a criterion level of ability or domain of content" (p.74). With reference to 'authenticity', Caroll (as cited in Bachman, 1990) suggests that alternative assessment is authentic as it makes reference to real life performance that is to say 'normal communication situation' and its functionality, which implies its 'total communicative effect'.

Huerta-Macias (1995) defines alternative assessment as "an alternative to standardized testing and all of the problems with such testing" (p. 8). While traditional forms of assessment such as standardized tests (e.g multiple choice, cloze tests) are claimed to mask what students really know or can do, alternative assessment enables students to be assessed on their demonstration of what they can do. According to Garcia and Pearson (as cited in Huerta-Macias, 1995, p. 8), that alternative assessment procedures consists of " efforts that do not adhere to the traditional criteria and standardization ..., objectivity and machine scorability. Hence, alternative assessment is different from traditional testing in that, it uses real-world settings, focuses on processes as much as products, helps to determine strengths and weaknesses of students, demands higher-level thinking and problem-

solving skills and ensures that scoring is done through human judgement (Brown & Hudson, 1998).

Although alternative assessment seems to have satisfied the need in the field of applied linguistics for more meaningful ways of assessment, still there are concerns about the validity, reliability and objectivity of these assessment procedures (Brown & Hudson,1998; Ekbatani & Pierson, 2000). Proponents of alternative assessment express their concerns about the validity of alternative assessment techniques for making decisions about people's lives, like placement and certification (Brown & Hudson, 1998). They also argue that compared to multiple choice tests that are scored objectively, alternative assessments require teacher judgement for scoring and therefore the probability of subjectivity and disagreement with other teachers is likely to be higher. Hence, lack of objectivity may result in unreliability of these assessment procedures (O'Malley & Pierce, 1998).

However, about the issue of validity Huerta-Macias (1995) explains that alternative assessments look at actual performance on real life tasks, like participation, writing and self-editing, and therefore the procedures are valid since all of them serve as concrete evidences of students' ability in using particular skills. Nevertheless, Brown & Hudson (1998) emphasise the importance of careful structuring, piloting, analysing and revising for improving the validity and reliability of these assessment procedures. They also highlight the importance of testers knowledge about "standard error of measurement and standards setting" (p. 656). When reliability is concerned, the idea advocated is that, a student's writing would demonstrate highly similar characteristics when graded over one week period by two different raters against the same holistic scale and will be rated either the same or

receive a similar score. Thus, to ensure that a score was based on actual student performance, O'Malley and Pierce (1998) recommend the use of a scoring rubric "that assigns a numerical value to the performance depending on the extent to which it meets pre-designed criteria" and rater training (p. 20-21). On the issue of objectivity, Huerta-Macias (1995) holds the idea that there is a human factor interfering both in standardized and in alternative tests. She adds that since we are humans, we all have biases this way or the other. As standardized tests are products of group of people who share the same biases, they are no more objectives than alternative assessment instruments.

This research studied self-assessment within the domain of writing and particularly in process writing approach. So, before going into self-assessment in more detail, some discussion of writing assessment in process writing and the use of alternative assessments in that field might be relevant at this juncture.

Alternative Assessments and Process Writing

The concept of process writing appeared in the field of ELT by the late 1970s. In process writing, the emphasis that was on the product with the traditional writing tasks shifted towards the writing process (Susser, 1994). The traditional approach to writing instruction emphasised correct usage and mechanics. It required students to study classical essay types (descriptive, narrative, argumentation), the rules that govern these types and also practice applying these rules. The success of students' writing was in turn measured by the students' ability to merge these rules into their writing. However, process writing approach promotes multiple drafting with teacher feedback between the drafts, provides meaningful writing opportunities on topics of interest or significance to the students, values content information, and

expression of personal thoughts and feelings more than grammar. This approach also raises students' awareness through the writing process (Grabe & Kaplan, 1996).

Students who practice writing through this approach become aware that writing is a process. According to Susser (1994) process writing pedagogies have two essential components: "awareness and intervention" (p. 34). This awareness is raised through instructional activities that help students to think thoroughly (brainstorm), organize their ideas before writing (outline), rethink and revise through multiple drafts during the writing process (Petrosky & Bartholomae,1986). These activities basically involve pre-writing, drafting, revising and editing. Intervention, which is claimed to be the other essential component of process writing, refers to the involvement and assistance of teacher during the writing process. Teachers in this process help students to think and organize their ideas before writing and also help them with the necessary revisions through the writing process.

Furthermore, unlike traditional writing approaches, process writing has also significant implications for writing assessment. The types of alternative assessments used in process writing approach have been specified as writing checklists, writing conferences, dialogue journals, learning logs, peer assessment and self-assessment (O'Malley & Pierce,1996). Regarding the assessment practices in the process writing approach O'Malley and Pierce (1996) also state that the assessment process is similar to and therefore need not be handled separately from classroom instruction. The way assessment is conducted in process writing approach imply changes in the writing assessment and instruction as is the case with all alternative types of assessment and innovative approaches to language instruction. These include changes in the

teachers' role in adapting assessment within instruction and involvement of student

in assessment and instruction.

Grabe and Kaplan (1996) argue that " alternative approaches to writing

assessment suggest relatively uncommon options for assessing student performance

that extend to the writing process for a specific essay" (p. 410). In relation to this

idea, they indicate self-assessment as a version of students' involvement in

assessment. They declare that self-assessment procedures should be used for specific

essay assignments and suggest a method for this in which the students are asked to

identify the strengths and weaknesses of a recently completed essay and show their

improvements reflected in this task over the former writing tasks. Finally, based on

their self-assessments students are asked to grade their essays and then the teacher

tells how she/he would assess the essay in terms of improvement and grade. This

type of assessment is advised for either negotiating a grade or for reaching at a

consensus about the student progress. This type of assessment is also proposed to

form a small portion of the overall grade allocated for a paper.

As the example from Grabe and Kaplan (1996) shows, assessment in the

process writing approach is not summative but formative. That is to say the major

concern is on the process rather than the product. During this process the teacher

helps the students with constant feedback to improve their writing skills and provides

the students with the opportunity to edit and revise their work as a part of the writing

assessment process. It has been suggested that, particularly in this approach, the

students should also be exposed to the scoring criteria, which their writing will be

graded against, and the prompt that defines the task to be performed (O'Malley &

Pierce, 1996).

Self-Assessment as an Alternative Assessment

It is commonly agreed that self-assessment is an essential learning strategy for autonomous language learning which capacitates students to closely monitor their improvement and relate learning to their individual needs (Harris,1997). In this context, McNamara and Deane (1995) also define self-assessment as students' critical awareness of their language learning process. Critical awareness refers to the students capacity to judge their own language abilities or language performance (Brown, 1998). Nunan (1988) argues that one of the goals of learner-centered curriculum for learners is to develop " a critical-self consciousness of their own role as active agents within the learning process" (pp. 134-135). He asserts that self-assessment is a very effective and efficient means of developing learners' awareness to their own capacity and how they can use this capacity.

Nunan (1988) also suggests that " in a learner centered curriculum model both teachers and learners need to be involved in evaluation" (p. 116). About the involvement of students in their own assessment process, Le Blanc and Painchaud (1985) state that " being part of the complete learning cycle should imply being involved in the assessment process, since evaluation is now being recognised as a component in the educational process" (p. 73). Thus, self-assessment is a practice that encourages students' reflection on their own learning process and involvement in this process.

Dickinson (1987) states that there are different purposes for which self-assessment can be used and also different learner groups from various ages and language levels whose degrees of involvement in self-assessment change correspondingly, and so categorises self-assessment as formative and summative.

As a summative assessment, self-assessment is claimed to provide a realistic possibility for the involvement of learners in making decisions about whether or when to take an examination about which learner's choices are seriously restricted. As a formative assessment, self-assessment is considered to function as: placement, achievement, diagnostic and progress assessments.

Self-assessment tests used for placement purposes have elements of certification in which learners can administer a pre-designed test to themselves and decide their level of competence on a given scale. When used for achievement purposes, self-assessment helps learners to measure their achievement in terms of a particular course that they have been studying. On the other hand, self-assessment used for diagnostic purposes provides information about learners' strengths and weaknesses in the target language and finally, continuous progress testing yields useful information both to the instructor and the learner on improvement within the course (Dickinson, 1987).

One strategy suggested for developing students' self-assessment is 'goal setting' which has been proposed as an ideal strategy for leading students to become more independent learners through raising their awareness towards their strengths, instructional needs and ways to achieve those needs (Smolen, Newman, Wathen and Lee, 1995). To encourage and assist goal setting, they suggest that students must first "learn how to critically examine their own work and to judge it against some standard they understand " (p. 22). Thus, through this strategy the students should learn what is good about their work, what needs to be improved and should develop the responsibility to make decisions for their own learning and become autonomous learners. O'Malley and Pierce (1996) suggest that teachers should also learn how to

help their students with this process and thus, build in various self-assessment approaches and strategies into their instructional goals. They propose exposing students to examples of good work and introducing the standards against which they have been judged. It has also been suggested that following these practices students should be given the opportunities to apply the criteria to the assessment of a sample selected among their own work. This particular suggestion points to the rationale for the training offered in this research.

Research on Self-Assessment

Research conducted on self-assessment in 1960s and 1970s was mostly concerned with comparing self-assessments with predictors of academic fulfilment and grades. However, it was not until the late 1970s and the mid 1980s that the studies focused on the self-assessment of language abilities (Ekbatani & Pierson, 2000). Together with the search for meaningful ways to engage the learners in their own assessment process the number of studies carried out on the use of self-assessment for the assessment of language skills increased. However, possible problems of self-assessment were set as the comparative subjectivity of grading, the variation in the accuracy of the scores depending on skill levels and materials involved in evaluation, and the possible unreliability of the scores in high-stakes situations (Brown, 1998).

Blanche (as cited in Brown & Hudson, 1998) forewarned that self-assessment scores may also be affected by subjective errors owing to factors like past academic records, career goals and expectations, and lack of training. He suggested that, if the students use scoring rubrics that consist of clear and precisely described criteria such subjective errors may be overcome.

Another concern raised about self-assessment has been whether learners possess the capacity and objectivity to view their own achievements (Dickinson, 1987). In a research study conducted by Le Blanc and Painchaud (1985), the researchers have argued that although self-assessment has been accepted for several years in the fields such as psychology, sociology and business, its use in second language teaching and learning is quite rare. Therefore their research sought to find whether it is because of distrust in the students' capacity to provide accurate information about their language skills or because of the possibility of using inappropriate practice of self-assessment. The research study aimed particularly at finding the answers to these questions: Do students have the ability to meaningfully evaluate their own performance? Does the type of instrument used affect that ability? With these purposes in mind a series of experiments were made leading to the use of self-assessment as a placement test. The research was carried out in the University of Ottawa, which is a bilingual university and first a sample of 200 students for both French and English as second languages were made to fill in a self-assessment questionnaire before taking the proficiency test in their second language. The total scores on this questionnaire were correlated with those on the proficiency test.

Then, at the second stage of the research two questions were asked on whether the content of the questionnaire or the variations in the formulation of statements for a given task could have effects on the results. To answer these questions two different questionnaires were given to students one of which included metalinguistic vocabulary and the other did not. It was found that the metalinguistic vocabulary had no significant effect unless the students are able to understand the language used in the questions. It was concluded that under the given conditions, self-assessment

functions as a valuable placement instrument since students find themselves responsible of their placement. Therefore it can be presumed that given appropriate, specific assessment tools learners should be able to properly rate their own abilities.

Another research study carried out by MacIntyre, Noels and Clement (1997) investigated perceived competence in an L2 as a function of actual competence and language anxiety. The participants of the study were 37 young adult Anglophone students who widely varied in terms of competence in French. The participants had completed scales of language anxiety and a can-do test, which altogether assessed their self-perceptions of competence on 26 French tasks. They then attempted each of these tasks. At the end of the study it was found out that perceived and actual L2 competences inter-correlated. However it was also found out that anxious students tended to underestimate and less anxious students tended to overestimate their competence.

Another study conducted by Ross (1998) suggests that self-assessment is a reliable alternative to formal second language assessment for placement and criterion referenced interpretations, although variation in self-assessment validity coefficients suggests difficulty in accurate interpretation. The research primarily studies the use of a formal meta-analysis conducted on 60 correlations reported in second language testing literature. Secondly, it covers a methodological analysis of the validity of a self-assessment instrument for which 236 EFL learners completed self-assessments of functional English skills derived from instructional materials and from general proficiency criteria. The learners' teachers also provided assessments of each of the learners. The criterion variable was an achievement test written to assess mastery of the completed course materials. Contrastive regression analysis was used and it

revealed differential validities for self-assessment as compared to teacher assessment depending on the level of learners' language skills. According to the results of the study it is seen that learners will be more accurate in the self-assessment process if the criterion variable illustrates achievement of functional 'can do' skills. Also, refreshing students' episodic/situational memory of using particular skills in the classroom experience would increase the accuracy of self-assessment.

According to the results of the studies examined so far, some of the factors that may affect self-assessment ratings are identified as: the level of proficiency of the students, their degree of anxiety, the type of self-assessment task, the particular instrument used for self-assessment purposes and lack of training.

Self-Assessment in Writing

Self-assessment in writing has been claimed to encourage the type of reflection that helps the learner to develop writer autonomy. Dickinson (1993) gives her own definition of autonomy " as an attitude to language learning which may not necessarily have many external, observable features" (p. 330) and lists the five main characteristics of autonomous learners as having the ability to identify what's been taught, determine their own learning objectives, decide and apply relevant learning strategies, eliminate the strategies that are not appropriate for them and, lastly, monitor their own learning.

O'Malley & Pierce (1996) argued that self-assessment makes the writers think about the purpose in writing and demonstrate the knowledge they have and how they can use it. They recommended four different methods for self-assessment in process writing. These are dialogue journals, learning logs, self-assessment of interests and checklists. In dialogue journals, students write on topics of their choice and address

their writing to their teacher, then teacher writes back modeling appropriate use of language. In learning logs, students make entries in the last five minutes of each lesson hour and try responding to questions like what they learned that hour, what was difficult to understand and what they need to do for better understanding. Surveys of interest and awareness are especially useful for teachers to learn about students' attitude to writing and monitor their writing improvement. Writing checklists provide the students with the opportunity to check their own writing with respect to the criteria contained in the scoring rubrics. For self-assessment of writing skills, Harris (1997) suggests that the criteria can either be outlined by the teacher or discussed with the whole class before each activity providing it to become an internal part of the writing process and the students may use the criteria as a checklist to guide their improvement of the particular skill. The final assessment can be compared to that of other students and the teacher's assessment. All these different types of self-assessment in writing have one common feature, which is interaction with instruction (O'Malley & Pierce, 1996).

Cohen (1994) defines writing assessment as " a complex interaction among three sets of factors: the knowledge that the test maker has about how to construct the task, the knowledge that the test takers have about how to do the task, and the knowledge that the test raters have about how to assess the task " (p. 307-308). Some suggestions to achieve the highest level of interaction between self-assessments of writing and instruction are to chose tasks that are appropriate for the learners, to chose rubrics that learners can make use of, to share these rubrics with the learners, to identify papers from different grade levels as models for the learners, to focus not only on what the learners write but also on how they write.

Also allocating time and feedback whenever needed by the learners, introducing self-assessment progressively through involving the learners in the process of the assessment of their own writing and modeling the editing process against the rubric being used and finally discussing their writing with the learners are considered as useful practices (O'Malley & Pierce, 1996).

Writing assessment has also been defined by Ferris and Hedgcock (1998) as " a formative and inherently pedagogical endeavour " (p. 227), which is closely related to other instructional processes. This research study will make use of a combination of the dimensions introduced in these definitions of writing assessment. With regard to this definition the current study employed set of standards for good writing; namely the writing criteria as an instructional tool for the self-assessment training the learners received, so it has a pedagogical dimension. This research also sought answers for whether the students learned how to use the given instructions, whether they could apply them and in case they could, did it result in any improvement in their writing skills. In this sense it has a formative, developmental dimension.

Training the Learners on Self-Assessment of their Writing

As discussed earlier, some important problems of self-assessment as identified by the professionals in the field are: relative subjectivity of scoring, possible variations in scoring depending on the students' language skills and materials involved in evaluation, and the possible unreliability of the scores in high-stakes situations. Professionals like Janssen-van Dieten (1989) and Pierce, Swain and Hart (1993) also raised serious questions about the learners ability to assess their own performance to which Harris (1997) replied:

> While doubts about the reliability of self-assessment have been raised most of these have been where students received no training. Jannssen- van Dieten (1989:p. 44) explains poor correlations between self- assessment and test results by lack of training: 'poor results plead for the application of self-assessment, rather than against it'. In other studies (e.g Bachman and Palmer 1989; Blanche 1990) excellent correlation between self-assessment and tests or teacher assessment have been found (p. 18).

This statement points to the fact that when trained, students are capable of assessing themselves accurately and objectively. So, one solution proposed for these problems is training the learners on self-assessment.

According to Dickinson (1987) the ability to assess the efficiency of one's own performance is an essential skill that gives the learners the opportunity to carry on the process independently and helps them gain autonomy over their own process of learning and this is why learners need to be trained on self-assessment. Hence, given the training the learners will be able to monitor their own learning process, develop the ability to identify their strengths and weaknesses, work towards solution of the problems they have and by this way gain control over their learning process (Dickinson, 1993).

One method that has been suggested for training the learners to self-assess their own writing is to use the grading criteria. Training the learners on the grading criteria has been considered as a means to introduce self-assessment gradually (Ferris and Hedgcock, 1998; O'Malley and Pierce, 1996). Harris (1997) states that with self-assessment of productive skills it is important to establish clear criteria for students to use when they evaluate their own performance. Research done on training of self-assessment of writing through the use of grading criteria has provided support for the above ideas.

Arter (1994) conducted a study to investigate the impact of training students to be self-assessors of their writing. The research was conducted with the teachers working with the Northwest Regional Education Laboratory who received training by this institution on a six-trait analytical scoring rubric used for assessing student writing. The six traits were ideas, organisation, voice, word choice, sentence fluency and conventions. The teachers were trained for 4 years to teach their students how to use the rubric to self-assess their writing. The purpose of the study was to explore the usefulness of this practice. The study was carried out with 67 students in the treatment group and 65 students in the control group who were 5th graders in six classes. Students had a pre-test before they started receiving the particular training which demonstrated similar results for both groups. After they received the training it was observed that the students in the treatment group performed better on most of the traits whereas the students in the control group were only able to improve on two traits.

Another study conducted by Hindman (1994) sought to identify the differences between the students' and the teachers' perceptions of qualities of good writing. In fact, the purpose was to help with the integration of the students in their own evaluation process. In this research graders of the Freshman Placement Exam had received training during which the criteria that the teachers use to evaluate student writing and how its use affects students' grades had been discussed. Also, involvement of the students in a discussion about the criteria was observed to be useful. After having the students in a basic writing class score some sample papers it was observed that the students were confused about the expectations of the teacher related to a trait in the criteria and were underestimating the significance of some

other traits like organization and style. This research again reveals the importance of teaching students how to use the writing criteria to be able evaluate and improve their writing skills.

Another study conducted by Ross, Rolheiser and Hogaboam-Gray (1998) searched for the effects of self-evaluation training on narrative writing. 148 students in 4 to 6 classrooms (15 grade) were trained for 8 weeks on how to assess their papers. Later their self-assessments were compared with the self-assessments of 148 students in the control group who received no training. As a result of the comparison of the self-assessment results of the two groups it was found that students in the treatment group were more accurate in their self-assessments than the students in the control group. It was also observed that the performance of the students in the treatment group was relatively better than the students in the control group and even those students in the treatment group who had poor writing skills had improved. The results of the treatment were ascribed to the effects of joint criteria development and its use, which was supposed to have increased the meaningfulness of self-assessment practice for the students.

Other than these recent studies in the field, there has also been other research on training the students to develop the ability to use the writing criteria for assessing and rating their writing tasks carried out back in 1970s and 1980s. One example is the research conducted by Clifford (as cited in Hillocks, 1986), in which sets of criteria were used to help students to assess their own writing. In his study the teachers were made to use what is named as " invariant sequence" for 13 compositions. The sequence begins with a structured assignment then goes on with teacher-directed oral brainstorming to explore different views, following a ten minute writing assignment

with response to the assignment disregarding mechanics and ends with discussion of ideas produced during this period in small-groups. In these discussions students were given feedback sheets to respond to each other's ideas about content and give suggestions for improving their writing. Following this stage the essays were returned to the writers and the groups exchanged their work for evaluation. Students then again were asked to evaluate papers of their friends and were given an evaluation sheet for this purpose on which they wrote about the weakest and the strongest parts of the paper. Each student read six papers other then their own paper applying the criteria and made suggestions for revision. They spent 28 of 35 hours during the semester. The experimental/control effect size of this research was reported as .61 and the pre-post effect size was reported as 1.12 for the experimental groups.

So far, related recent and past studies which all employed grading criteria for the purpose of training the students to self-assess their own writing have been reviewed. Looking at the results of the studies, one common suggestion has been the use of clear and understandable grading criteria for effective training of the learners which is claimed to yield positive results. Thus, the purpose of this study is to see whether training the learners to self-assess their writing, through the use of the grading criteria, will have an effect on the quality of their assessment and on the improvement of their writing skills as suggested by the literature.

CHAPTER THREE: METHODOLOGY

Introduction

This research investigated the effect of training students to self-assess their own writing. More specifically, the research studied the effect of trained self-assessment versus untrained self-assessment when students use the same criteria. This was done through the comparison of the self-assessment results gathered from one control and one treatment group. The students in the treatment group received training on how to use the criteria for self-assessment purposes whereas the students in the control group self-assessed their writings using the same criteria without any training. To find an answer to the research question, the students' self-assessments were compared with the assessment of their teacher. The closer the assessment of the students to their teacher's assessment, the more effective the training was considered. This study also explored whether the training on students' self-assessment of their writing had any effect on the improvement of their writing skills. This research question will be addressed through comparison of the assessment results of the external raters (the instructor and the researcher) for each draft of the students' writings in the treatment and the control groups.

Participants

The participants in this research were 25 Bilkent University, First Year English students. They were 101 Freshman Engineering students who received content based instruction courses and practiced reading-based process writing. Two classes participated in the study. In the first class there were 13 students (all male) and in the second class there were 15 students (12 males and 3 females). Students in

both classes were irregular students who received English 101 in the second semester. Irregular students are those who can attend First Year English classes only after studying English in Bilkent University School of English Language (BUSEL) until they perform adequately in Certificate of Proficiency in English Exam (COPE). So, those students who study English in BUSEL and start the courses later than their regular semester are called irregular students.

The treatment and the control groups were selected randomly from these two classes through a raffle. As a result, the students in the science class were selected as the treatment group and the students in the engineering class as the control group. In the treatment group there were 15 students, one of whom dropped the course before the training had started and another student did not participate in the research, so the number of the students in this group dropped to 13. In the control group there were 13 students one of whom dropped the course and did not participate in the study, so the number of the students in the control group dropped to 12.

The levels of the students were upper-intermediate. In order to ensure that the students in the two groups were of similar language levels, their COPE proficiency scores were collected from the First Year English Department. Also their End of Course Assessment (ECA) writing scores were collected from BUSEL to see whether there is a significant difference, between the writing skills of the students in the two groups, that might effect the results of the study. With reference to the data, it was seen that in both groups there were 10 students who received B in COPE and the rest had received C. Also the with regard to their ECA writing scores the students in both groups were fairly similar. In the treatment group, out of 13 students 6 students couldn't pass the September COPE and the ECA writing scores of the other

7 students varied between 7 to 18 over 20 (one 8, two 13, one 14, one 15, one 18). In the control group, out of 12 students 7 students failed the September COPE and the ECA writing scores of the other 5 students ranged between 7 to 15 over 20 (one 7, three 14, one 15). So the mean ECA writing score for the treatment group was 13.16 and the mean ECA writing score for the control group was 12.8, both of which correspond to 13.

The participants' ages ranged between 18-20 and in both groups they had studied English in BUSEL. In the experimental group, out of 13 students 2 of them had studied English for 3 semesters in BUSEL and the rest had studied English for 1 semester. In the control group, out of 12 students 1 had studied English for 4 semesters in BUSEL and 11 of the participants had studied English for 1 semester.

The reason for choosing irregular 101 English students was primarily because they were practicing essay writing, and secondly because of the time constraints. In the first academic semester, students in English 101 classes practice essay writing, and in the second academic semester, students in English 102 classes produce research papers and study research skills. As one of the questions in this research was whether training had any effect on the improvement of the students' writing skills, 101 English students studying essay writing through process writing approach were preferred. Then it would be possible to compare and see the improvement between the drafts. It was also because of the time constraints. Since the research had started in the second semester there was no other choice left but to work with irregular students.

Research Design

This research study was conducted at Bilkent University, First Year English Program to explore the possible effects of training students to self-assess their own writing, using the course criteria. In this study, the data were collected through the participation of one First Year English instructor and two 101 First Year English classes. Data were collected in four different phases. The students in both groups were first given a background questionnaire. Then the students in the treatment group were given training on how to self-assess using the criteria, then they self-assessed and marked their own writing. The students in the control group also self-assessed using and rated their writing against the same criteria, but without training. After that the papers of the students in both groups were assessed and graded by two external raters (their teacher and the researcher). In the last stage, the students in the treatment group were given a questionnaire asking about their attitudes toward self-assessment.

For training, the generic writing criteria that were used in the assessment of ENG 101 for Engineering and Science students' compositions in 2000-2001 Fall Semester (see Appendix A) were used. The study basically dealt with training the students to be able to interpret and apply the writing criteria on their own with a view to meet the requirements of the given assignment while producing and assessing their writing. For this purpose, the suggestions made by Hillocks on different ways to teach criteria for better writing were considered. Hillocks (1986) proposes to (1) have the students study model pieces of writing to gain the ability to identify qualities of good writing (2) use scales or sets of criteria and applying them to sample models of writing and other writing (3) review and consideration of teacher feedback.Hillocks (1986) also emphasises that the first two methods are for pre-writing process aimed

at teaching the learners how to use the criteria before they start the writing process and the following method is post-writing instruction in which students learn the criteria as a result of their writing process, to make use of in their future performances.

In this research both pre-writing and post-writing methods were employed. The students were asked both to apply the writing criteria to their own papers and also make use of the instructor's feedback for self-assessment within the framework of the particular training that was offered in the research study.

<div align="center">Materials</div>

<u>Writing Criteria</u>

The criteria used in this research were generic writing criteria used in assessing writings of ENG 101 for Engineering and Science students in 2000-2001 Fall Semester. It is a three trait analytic rubric against which the students' writings are assessed. The first trait is Content which covers 50% of their writing, the second trait is Organization which covers 30% of students' writing and the third one is Vocabulary/Sentence structure/Mechanics which covers 20% of the total grade that the students receive. In all these three parts, the criteria for scoring are defined for all different levels of performance from A to F (see Appendix A).

The percentage of each trait is determined according to their importance with regard to the course objectives. Since all writing is based on content reading, the students are expected to reflect in their writings the content information they gained throughout the course. So the content is of primary importance and that is why it is 50%. As the students are required to reflect the content knowledge they have gained throughout the course in their writing, in the assessment stage the reading texts are

referred to see whether they could provide the specific information or evidence of content knowledge.

For this study, the particular content that the students studied was " Family". The substantial material used in the course was *Sons and Lovers* by D. H. Lawrence. Although organization seems to be of secondary importance when compared with content, it is a crucial element of the writing process as 101 English classes are actually 'Writing and Composition' classes. Also, the instructors are allowed to specify the existing requirements in the writing criteria or to make modifications to adopt the criteria accordingly for different types of writing assignments. For this study, the criteria were modified for self-assessment purposes (see Appendix E). For the modification the subject of the criteria was changed from 'The student' to 'I', check-boxes were placed beside each descriptive line for the students to tick in, upon their achievement of the particular requirement defined in that line. Also, the criteria were specified based on the requirements set with the prompt (e.g "sufficient references" was replaced with "at least three references").

The writing task used was argumentative essay and the way they practiced writing was through process writing. Argumentative essay was used mainly because the students were practicing argumentative writing during the research period.

Model essay

The model essay used was the final draft of an argumentative essay written by a regular English 101 Engineering student who took the course from the same instructor (see Appendix B). In order to be able to use the model essay for training purposes, the owner of the paper was asked to give permission and signed a consent form (see Appendix C)declaring that her paper could be used as the model paper.

The model paper used in this training was chosen based on the suggestions made by O'Malley and Pierce(1996), according to which " One of the ways to communicate students to what good writing looks like is to select benchmark papers,(i.e., papers that you have rated high on the components of your scoring rubric). Share the papers with students as models they can emulate" (p.159). As judged by the instructor and the researcher rater, it was a well organized argumentative essay, the content was handled effectively with clear references and examples, language was clear and accurate.

There were two other reasons for why the particular essay was chosen to model the assessment procedure. First, the paper was written upon a similar task with that of the participants' in this research study. In addition, the model essay was performed with reference to the same prompt with the exception of content. The content of the model paper was based on "Fantastic Literature" and the substantial material that had been studied was *Hobbit* by J.R. Tolkien.(see Appendix D).

Questionnaires

The first questionnaire that the students received was a background questionnaire in which students were asked to provide information about their age, department, language level and their writing experience in general (see Appendix G).

The second questionnaire was given at the end of the study. This questionnaire was given to identify the attitudes of the students toward self-assessment of their writing. In this questionnaire, the students were asked questions about the effectiveness of the training, the efficiency of the components of training like the model essay, generic writing criteria, self-assessment criteria and their attitudes about self-assessment. (see Appendix H).

Prompt

Both the prompt used for the model essay and the prompt used for the argumentative essay (see Appendix F) were based on the format adopted by the Professional Development Person, in the First Year English Program at Bilkent University. This format had been used by the instructor of the participants, so special permission was asked from the PD Person.

Piloting

The study was piloted in the 8th and the 9th weeks of the Spring Semester with three management students who were also irregular students receiving 101 English in the second semester and studying argumentative essay writing. The students in that class were given information about the study and asked whether they would like to participate in the pilot study. The three students (2 male, 1 female) were then chosen on a voluntary basis. In this first pilot session, the students' training which was planned to last one class hour lasted two class hours. This could be due to the fact that the students were allowed to interrupt during the training and ask questions or comment on the model essay and the discussion after the training (on the scores) lasted longer than expected. It could also be because the researchers' demonstration of the grading of the model paper was rather confusing as all the three traits were analysed in an almost holistic fashion against the criteria. Also during this session one of the students tended to speak in Turkish most of the time and continuously made early comments without listening to the end of the analysis and interrupting the researcher. Also these students tended not to refer to the criteria for grading and they all gave letter grades to the model paper. At the end of the first session the students were asked to meet again for a second training.

In the second session, the students were primarily reminded of the stages in the previous session of training. Also they were introduced and given information about the self-assessment criteria that were used in the second session of the training (subject of the criteria was changed to " I " and check boxes were placed beside each descriptive line). In this session, the students were only monitored and provided with help when they needed.

The piloting yielded useful results based on which some important decisions were taken and the method of demonstration was revised with regard to the problems encountered during these sessions. The decisions taken after the pilot sessions were as follows:

a) give more and detailed information about the criteria

b) explain in detail what each trait is all about (content, organization, vocabulary /sentence structure/mechanics)

c) define how they will decide the quality line to place their paper for each trait (from A to F)

d) remind the students of the prompt and the points they should be careful about (the criteria refers to the prompt in the content trait – task fulfilment)

e) review and revise the instructions given during the training to prevent any misunderstandings, confusions , discussions and to provide a smooth flow of the training

f) tell the students to

• listen to the explanations attentively and in silence

• not to comment on anything or ask questions before the training finishes

• refer to the criteria for grading

- give score grades but not letter grades

g) grade the paper first for the content, secondly for the organization and lastly for the vocabulary/sentence structure/mechanics. Make sure all three traits won't be dealt with at once. Handle the demonstration separately for each trait.

Training Procedures

During the first stage of the research, the students in both groups were given a background questionnaire in which they were asked to give information about their age, gender, department, language level and writing experience.

Following the background questionnaire, students in the treatment group were trained to self-assess their own writing. The students in the control group self-assessed their papers too but they did not receive training. Instead they continued with their usual classroom instruction.

The training for the treatment group was planned in the following four phases.

1) Introduction and Orientation to Application

2) Demonstration

3) Discussion

4) Practice

In the introduction phase, the students in the treatment group were first given information about the subject of the research and it's purpose. Also, in this phase, points like the use of self-assessment and how they may benefit from this training were explained. Following the warm up, the students were given the model essay that was going to be used for training purposes and oriented to its' assessment. First, the students were given the necessary information about the content of the model paper and criteria and then they were asked to read attentively and to score the paper

referring to the criteria. The criteria they used were the generic writing criteria that had not been modified then because in this stage the aim was just to see whether they could use the criteria or not. Also, during this first phase, they did not assess their own papers but a model paper. For the later stages of training the set of criteria were adapted for self-assessment purposes.

In the demonstration phase, after the students assessed and scored the model paper, the trainer (the researcher) on the overhead projector assessed the model paper. By so doing, the students were exposed to the evaluation and the scoring of the same model paper, and in a way observed how the mental process of their instructor might work while assessing the paper. The demonstration stage was followed by the discussion stage.

In the discussion phase, the result of the trainer's assessment and those of the students' assessment were compared. Unlike the piloting results, this time, the trainer's assessment and the students' assessments were quite similar. These results may be due to the modifications made and better-organized training sessions with clear instructions and directives.

In the practice phase, the students in both groups assessed the first, the second and the third drafts of their argumentative essay assignment on the due dates of each draft. Before the students self-assessed each draft of their essays, the above training procedure was repeated for the treatment group but in a shorter period of time. As mentioned before, the criteria were modified for this phase, for the students to feel comfortable while self-assessing. The major modifications were made through changing the subject of the criteria from " The student has " to " I have", placing check-boxes beside each definition and specifying the definitions. In this hour, the

students in both the control and the treatment group were asked to assess and score their own papers with reference to criteria that was modified for self-assessment purposes. When they were self-assessing their papers, the students were only monitored with no intervention made. The papers were then read and scored by two independent raters (their instructor and the researcher). It is important to note that, the students' essays were read and scored blindly for reliability purposes.

In the last stage of the training, the students in the treatment group were also administered an attitude questionnaire that consisted of questions on how they felt about the training and the practice of self-assessment.

CHAPTER 4: DATA ANALYSIS

Introduction

The purpose of this study was to investigate the effect of training the students on the self-assessment of their writing. This study sought to answer two research questions: (1) Does training students to self-assess their own writing, using the course criteria, affect the quality of self-assessment? (2) Does training students to self-assess their own writing, using the course criteria, improve their writing skills? The independent variable was the training offered to the students and the dependent variable was the students' self-assessment of their writing. In this chapter, the answers to the two research questions were provided by looking for the possible changes in the dependent variable caused by the independent variable. This was done by the use of several different statistical procedures. The data collection instruments used for this purpose were a background questionnaire, three drafts of students' essays as assessed both by the external raters (the instructor and the researcher), students' self-assessment records, and the attitude questionnaire given at the end of the training process.

The statistical procedures used to analyse the data are: Pearson Correlation, t-test and comparison of mean scores. Also, an inter-rater reliability was used to check the degree of reliability between the assessments of the two raters. Pearson Correlation and t-test were computed to answer the first research question and thus, to see whether the training caused the treatment group to perform significantly better in the self-assessment of their writing than the control group. Comparison of the mean scores and again t-test were used to answer the second research question. Comparison of the mean scores was used to determine whether there is a difference

between the raters' assessments of the three drafts of the students' writings in the treatment and the control groups. T-test was used to find out whether the training caused an improvement in the writing skills of the students in the treatment group.

In the following sections, the data analysed through the use of the above mentioned statistical procedures are presented.

Data Analysis

Reliability

In this study, the students' writings were scored by two independent raters (their instructor and the researcher) using the generic writing criteria. In order to establish reliability of the scorings used, the inter-rater reliability of the two raters was checked. Inter-rater reliability was determined by finding out the relationship between the two raters' scorings. As a result, three reliability coefficients were obtained, each for each of the three drafts of the essay. The reliability of the ratings for the first drafts of the students' essays was .65, for the second draft it was .51 and for the final draft it was .71. As can be noted, the reliability of the first two scorings were not very high, though the third one seems to be high enough to say that the scoring of the third draft was almost reliable. However, in order to have a more valid index of the students' writing skills, the scores of the two raters were combined.

The Effect of Training on Students' Self-Assessment

Pearson Product Moment Correlation was computed to see the extent to which the students' ratings were similar to or different from that of the two raters' combined ratings (hereafter the external ratings) when they are trained as opposed to when they are not trained. This analysis was intended to provide the answer to the first research question (the effect of training students to self-assess their own writing,

using the course criteria, on the quality of their self-assessment). In this statistical

measurement procedure, the difference between the external raters' and the students'

scorings for each of the three drafts of the essay were compared. The closer the

students' assessment to that of the external raters, the more effective the training was

considered. The results of the analyses are illustrated in Table 1.

Table 1

The Pearson Correlation Results on the Difference Between the External Raters'
Scoring and the Students' Self-Assessment Scores for the First, Second and the Third
Drafts of the Essay

Groups	Ratings	EFD	ESD	ETD	SFD	SSD	STD
	EFD						
Treat.	ESD	.861**					
N=13	ETD	.839**	.979**				
	SFD	-.139	-.007	-.041			
	SSD	.210	.249	.200	.685*		
	STD	.323	.454	.401	.574*	.835*	
Contr.	EFD	1.000					
N=12	ESD	.831*					
	ETD	.809**	.919**				
	SFD	.418	.663**	.578**			
	SSD	.474	.741**	.587*	.702*		
	STD	.354	.625*	.487	.547	.860**	

Note. EFD= external raters' combined score for the first draft, ESD= external raters' combined score
for the second draft, ETD= external raters' combined score for the third draft
SFD= students' self-assessment scores for the first draft, SSD= students' self-assessment scores
for the second draft, STD= students' self-assessment scores for the third draft
** Correlation is significant at the 0.01 level(2-tailed).
* Correlation is significant at the 0.05 level (2-tailed).

As the table shows, the correlation of the scores given by external raters and the

students in the treatment group was -.139 for the first draft of the essay, for the

second draft it was .249 and for the third draft it was .401. The findings show that

there is a constant and consistent increase in the scores of the students in the

treatment group. This consistency in pattern can not be observed for the control

group. The correlation of the scorings by the external raters and the students in the

control group were .418 for the first draft, .741 for the second draft and .487 for the final draft. However, still none of the increases were significant at .05 level.

Another important result was the high and significant correlation of the external raters in their ratings of the second (.861) and the third drafts (.839) of the essay. Considering the initial differences in the scorings of the two external raters, this results indicate that, when combined, the external raters' ratings within and across the three drafts of the essay is both more consistent and correlates higher. In fact, the inter-rater reliability of the two raters increased.

In sum, it can be concluded from the analysis of the data that the training seems to have been helpful in making the scoring of the students in the treatment group get closer to the external raters' scoring. It was observed that the scorings of the students in the treatment group got closer to the scorings of the external raters over the course of the research. In other words, training added to the validity of the self-assessment to increase but not much. The consistent increase in the self-assessment scores of the students in the treatment group can be explained in relation to two possible effects of the training: developing learner autonomy and explicit teaching. On the other hand, the results of the analysis for the control group demonstrated an inconsistent correlation. Although there seems to be a significant increase in their scoring in the second draft of the essay, in the third draft a distinct decrease is observed.

In order to see whether there is any difference between the raters' and the students' self-assessment when they have received training versus when they have not received training t-test was run. It was hypothesized that the smaller the difference gets between the self-assessment scores of the students and the external

raters' rating over the course of the research, the more valid the self-assessment becomes. In other words the researcher expected to see that the degree of difference between the external raters and the students in the treatment group as opposed to the control group would decrease as a function of the training. With this analysis, the researcher hoped to show that the difference has got smaller when the students have received training. For this purpose, Paired Samples Statistics T-Test was computed between the raters' and the students' assessments for three drafts of their writing. The results of these analyses are demonstrated in Table 2.

Table 2

The T-Test Results on the Difference Between the Scores of the External Raters and the Students' Self-Assessment Scores for the First, Second and the Third Drafts of the Essay

Groups	Ratings	Mean	SD	Std. Error Mean	t-value	Sig (2- tailed)
	EFD	74.84	14.47	4.01	-3.008	.009
Treat.	SFD	88.30	4.42	1.22		
N=13	ESD	82.76	10.78	2.99	-2.469	.030
	SSD	89.92	2.98	.82		
	ETD	92.84	3.33	.92	2.690	.020
	STD	86.61	9.10	2.52		
Contr.	EFD	77.58	11.28	3.25	-2.896	.501
N=12	SFD	86.41	7.29	2.10		
	ESD	85.29	7.39	2.13	-4.173	.002
	SSD	91.33	6.15	1.77		
	ETD	89.66	7.54	2.17	2.313	.041
	STD	94.41	6.35	1.83		

Note. EFD= external raters' combined score for the first draft, ESD= external raters' combined score for the second draft, ETD= external raters' combined score for the third draft
SFD= students' self-assessment scores for the first draft, SSD= students' self-assessment scores for the second draft, STD= students' self-assessment scores for the third draft

The results of the t-test indicated that in the treatment group the students rated themselves higher in the first two drafts of their writings and lower in the final draft. On the other hand, the students in the control group seem to have scored themselves higher than the raters in all three drafts. The reason for why the students in the treatment group self-assessed lower in the final draft of their writings may be due to

the awareness they raised toward their own performance after the training and hence

they seem to be more cautious in their scoring

To see, whether there was any change in the performance of the groups, mean

difference were calculated and compared across groups. The results of this

procedure are presented in Tables 3.

Table 3

The Paired Mean Difference Between the External Raters' and the Students' Scoring
 of the First, Second and the Third Drafts of the Essay

Ratings	Mean (Treatment)	Mean(Control)
EFD	-13.46	-8.83
SFD		
ESD	-7.15	-6.04
SSD		
ETD	6.23	4.75
STD		

Note. EFD= external raters' combined score for the first draft, ESD= external raters' combined score
for the second draft, ETD= external raters' combined score for the third draft
SFD= students' self-assessment scores for the first draft, SSD= students' self-assessment scores
for the second draft, STD= students' self-assessment scores for the third draft

As demonstrated in Table 3 there was a difference of 7.3 point between the first

draft (−13.5) and the final drafts of their writings (6.2) for the treatment group.

Whereas, there was a difference of 4.1 point between the first (8.8) and the final

drafts of their writings (4.7) for the control group. Hence, in the light of this

evidence, it can be argued that the difference between the teachers' ratings and the

students' self-assessment scores has become smaller for the treatment group as a

function of the training.

The Effect of Training on Students' Writing Improvement

In the next stage of the data analysis, mean gain scores of the treatment and the

control group as assessed by the external raters were compared for each draft of the

essay in order to see whether there is a difference between the groups in terms of

writing improvement. It was hypothesized that a statistically significant increase in the scores of the treatment group would indicate that the training had contributed to the improvement of the students' writing skills. The results of the analysis are as shown in Table 4.

Table 4

The Comparison of Mean Scores of the External Raters for the First, Second and Third Drafts of the Essay

Group		FD	SD	TD
Treatment	Mean	74.84	82.76	86.61
	N	13	13	13
	SD	14.47	10.78	9.10
Control	Mean	77.58	85.29	89.66
	N	12	12	12
	SD	11.28	7.39	7.54
Sig.		.605	.506	.373

Note. FD: first draft, SD: second draft, TD: third draft
*$p < .05$

According to the results of the mean comparison, although the control group started with a higher mean (77.58) than that of the treatment group (74.84), the difference was not statistically significant.The performance of the two groups in terms of their writing skills were considered similar when they began writing. Their language levels were shown to be similar based on their COPE proficiency scores reported by BUSEL.

The accepted level of significance for this analysis was $p < .05$. whereas the significance level for the first draft was .605, for the second draft it was . 506, and for the final draft it was .373, all of which were not significant. Although both groups were observed to have improved significantly in terms of their writing skills throught the writing process, the comparison of the two groups' means showed that there was no statistically significant difference between the two groups in their first, second

and third drafts. As a result, the training seem not to have a big effect on the students' writing skills.

In addition to mean comparison, Paired Sample Statistics was also computed to see whether the training had an effect on the students' improvement of their writing through the first, second and third drafts of the writing as assessed by the raters and as assessed by the students. The results of the data analysis are shown in Table 5 and 6.

Table 5

The Results of Paired Sample Test on the Improvement of the Students' Writing Based on the External Raters' Assessment

Groups		Mean	N	SD	t-value	df	Sig. (2-tailed)
Treatment	EFD	74.84	13	14.47	-3.783	12	.003
	ESD	82.76	13	10.78			
	ESD	82.76	13	10.78	-5.299	12	.000
	ETD	86.61	13	9.10			
Control	EFD	77.58	12	11.28	-4.056	11	.002
	ESD	85.29	12	7.39			
	ESD	85.29	12	7.39	-5.038	11	.000
	ETD	89.66	12	7.54			

Note. EFD= external raters' combined score for the first draft, ESD= external raters' combined score for the second draft, ETD= external raters' combined score for the third draft
SFD= students' self-assessment scores for the first draft, SSD= students' self-assessment scores for the second draft, STD= students' self-assessment scores for the third draft

The results indicate that difference between the two groups was not significant for successive drafts as assessed by the external raters.

Table 6

The Results of Paired Sample T-Test on the Improvement of the Students' Writing
Based on the Students' Self-Assessment

Groups		Mean	N	SD	t-value	df	Sig. (2-tailed)
Treatment	SFD	88.30	13	4.42	-1.749	12	.106
	SSD	89.92	13	2.98			
	SSD	89.92	13	2.98	-5.707	12	.000
	STD	92.84	13	3.33			
Control	SFD	86.41	12	7.29	-3.214	11	.008
	SSD	91.33	12	6.15			
	SSD	91.33	12	6.15	-3.222	11	.008
	STD	94.41	12	6.35			

Note. EFD= external raters' combined score for the first draft, ESD= external raters' combined score
for the second draft, ETD= external raters' combined score for the third draft
SFD= students' self-assessment scores for the first draft, SSD= students' self-assessment scores
for the second draft, STD= students' self-assessment scores for the third draft

As assessed by the students, again a similar result was obtained. Based on their

own self-assessment, the improvement of the writing skills of students in the

treatment group between the first and the second draft was from 88.30 to 89.92 and

between the second and the third draft it was from 89.92 to 92.84. Again based on

their self-assessment results, the improvement of the writing skills of students in the

control group between the first and the second draft was from 86.41 to 91.33 and

between the second and the third draft it was from 91.33 to 94.41. These data also

proved that the students improved irrespective of the training.

The results obtained for the second research question can also be interpreted as

a natural consequence of the writing process. Since in process writing the aim is to

improve students' writing by giving them the opportunity to write and revise their

essays in several drafts, usually students' grades improve.

Attitude Questionnaire Results

The attitude questionnaire was only administered to the students in the

treatment group (13 students) as it consisted of questions related to the given training

on self-assessment of their writing. The data collected with the attitude questionnaire

was analysed in terms of the frequency of the answers given by the students to each

question and their percentages. The percentages of the students' answers are given

beside each choice offered to the students for each question. In the analysis of the

data the statistically valid answers given by the students to each question were

considered. The results of the analysis are as demonstrated below.

Question1   Do you think the training given to you on the self-assessment of your

papers was effective?

a) Yes (76.9 %)   b) No (23.1%)

As can be seen, majority of the students thought that the training given to

them on self-assessment of their writing was effective.

Question 2   If your answer to the first question is Yes, then please circle the most

appropriate reasons for why you think the training was effective? Check all that

apply.

  a) I developed the habit of using the criteria effectively. (15.4 %)

  b) I can now see the strengths and weaknesses of my paper. (53.8 %)

  c) I can now predict the points on which my teacher will give feedback. (23.1 %)

  d) I can now make realistic guesses about the grade I will receive. (7.7 %)

  e) I believe that I learned how to self-assess. (23.1 %)

For those students who thought the training was effective, the primary reason was that they could see the strengths and weaknesses of their papers. The secondary reasons are being able to predict the points on which their teacher will give feedback and having learned how to self-assess. Lastly, the students seem to have thought that the training have been relatively less effective in helping them develop the habit of using the criteria effectively and making realistic guesses about their grades. Regarding the answers of the students to this question, they seemed to have considered the effect of training in relation to the extent to which it met their individual needs as learners. Thus, the results point out to a raise of awareness in the students towards their own learning process.

<u>Question 3</u>   If your answer to the first question is No, then please circle the most appropriate reasons for you out of the below given choices? Check all that apply.

   a) I still can't use the criteria effectively.(7.7 %)

   b) I still can't see the problems in my paper. (100.0 %)

   c) I can't trust in my own assessment. (15.4 %)

   d) I can't guess the points that I will get feedback on from my teacher. (100.0 %)

   d) I don't think that I learned how to self-assess. (100.0 %)

For those students who thought the training was not effective, the most important reason was that they couldn't trust in their own assessment and the least important reason was that they still couldn't use the criteria effectively.

<u>Question 4</u> Would you like to go on with self-assessment by yourself in your future writing assignments?

a) Yes (69.2 %)   b) No (30.8 %)

Most of the students said that they would like to go on with self-assessment by themselves in their future assignments. This, in fact, is a positive result of the study.

Question 5   What do you think was the most useful part of the training on self-assessment of your own writing? Please circle the answers that best express your thoughts. Check all that apply.

   a)   learning how to use the criteria (7.7 %)

   b)   having control over the assessment of my paper (38.5 %)

   c)   learning how to self-assess (23.1 %)

   d)learning how to identify the problems in my paper and to revise them (61.5 %)

   d)   improving my writing (30.8 %)

For most of the students the most useful part of the training on self-assessment of their writing was learning how to identify the problems in their paper and how to revise them. They thought that the second most useful part of the training was to help them gain control over the assessment of their paper. Improving their writing and learning how to self-assess seemed to be relatively less important for the students. Finally, the students thought that the least useful part of the training was learning how to use the criteria. These results indicate two important outcomes. First, the students seem to have developed autonomy and felt more in control of their own learning process. Second, although they seem to be aware of the benefits of learning how to use the criteria for self-assessment purposes, they still seem to fail in understanding its' importance for the achievement of their individual goals as learners.

Question 6   Which essay was more difficult to assess for you?

a) the model essay used in the first session of the training. (7.7 %)

b) my own essay. (92.3 %)

For almost all of the students assessing their own essays was more difficult. This result may be an indicator of the interference of subjectivity.

Question 7   Which criteria were you more comfortable using?

a) generic writing criteria, the subject of which was " The student " (46.2 %)

b) self-assessment criteria, the subject of which was " I " (53.8 %)

More than half of the students felt more comfortable using the self-assessment criteria, the subject of which was "I". This result may be due to the simplifications made in the self-assessment criteria to offer the students with a more clear and precise set of criteria to work more comfortably. Also, close to half of the students felt more comfortable using the generic criteria. This result may be due to their familiarity with the generic writing criteria.

Question 8   What do you think were the positive (+) and the negative (-) aspects of the self-assessment criteria? Put a check in the box that best reflects your thoughts about each item on the left.

	Positive	Negative
a) the subject " I "	☐ (69.2 %)	☐(30.8 %)
b) the check boxes beside each descriptive line	☐ (92.3 %)	☐(7.7 %)
c) the categories of the criteria (content, organization, vocabulary/sentence structure/mechanics)	☐ (100.0 %)	☐(100.0%)
d) the language of the criteria	☐ (84.6 %)	☐(15.4 %)
e) the practicality of scoring	☐ (69.2 %)	☐(30.8 %)

For a significant number of the students the check boxes beside each descriptive line and the language of the criteria were positive aspects of the self-assessment criteria. Only a small number of the students thought that these were the negative features of the criteria. While more than half of the students considered the subject "I" and the practicality of the scoring as positive aspects of the criteria, still a comparably smaller percentage of the students thought that these aspects were negative. The students were neutral about the categories of the criteria and did not consider them either as positive or a negative aspect of the criteria. The results indicate that the check boxes may have made it easier for the students to monitor their achievement and progress. Also, simplicity and clarity of the language may have eased the practice of self-assessment process for the students.

Question 9   Which of the three traits (content, organization, mechanics) was more difficult to score?

   a) content (61.5 %)

   b) organization (30.8 %)

   c) mechanics (7.7 %)

For more than half of the students content trait was more difficult to score when compared to the other two traits, organization trait was less difficult to score and the least difficult trait to score was the mechanics. These results point out an important fact to be considered for the future training implications. The students need to be trained more on the content of the particular writing task.

Question 10   If you were to receive another training on the self-assessment of your writing skills, which of the following would you like to have?

   a) simpler criteria (23.1 %)

b) longer time for training (30.8 %)

c) more detailed instruction about scoring (46.2 %)

Plurality of the students wanted to have more detailed instruction about scoring, if they were to receive another training. Their second concern was to have longer time for training and lastly, a small percentage of the students wanted to have simpler criteria.

Overall it seems that the students thought the training was useful in general. It is also revealed in the results of the questionnaire that their attitudes towards self-assessment of their papers are in general positive. It seems as if the students were happier about learning how to identify the strengths and weaknesses of their papers and revising their writing. One other important conclusion that can be drawn out of the results is that subjectivity may interfere while they are assessing their own essays. This result was determined based upon the students' answers saying it was more difficult to assess their own essays than assessing the model essay. The possible interference of subjectivity can be lessened through use of specific and clear set of criteria on which more than half of the students agreed to have felt more comfortable using. Also, some further points to consider about the training seem to be the need to elaborate more on content evaluation, to leave longer time for training and to give more information on the scoring procedure.

CHAPTER 5: CONCLUSION

Overview of the Study

This study aimed to discover (1) the effect of training students to self assess

their own writing, using the course criteria, on the quality of their self-assessment,

and (2) the effect of training students to self-assess their own writing, using the

course criteria, on the improvement of their writing skills. The focus of this study

was to train the students on self-assessment of their writing skills to enable them to

be able to monitor their own improvement, to raise their awareness towards their

strengths and weaknesses as a writer and thus, to become more autonomous learners.

The study was conducted at Bilkent University, First Year English Program.

The participants of the study were 25 irregular Freshman Engineering and Science

students and their instructor. The groups were decided randomly and the participants

volunteered to take part in the study. There were 13 students in the treatment and 12

students in the control group. Students in the treatment group had a training of four

sessions given by the researcher during the students' regular class hour. The first

session of training lasted 60 minutes. The other sessions of training were shorter and

were in the form of a brief review and practice of all that had been taught in the first

session of training. Meanwhile, the students in the control group also practiced self-

assessment identically but without any training and went on with their usual

instruction.

The materials used for training purposes were FYE 2000-2001 generic

writing criteria, self-assessment criteria and a model essay. The generic writing

criteria were the criteria that their instructor used to grade their writings and had

given to the students, so the students were familiar with this criteria. The self-

assessment criteria was the modified version of the generic criteria for self-assessment purposes. This criteria was changed by the researcher upon the agreed modifications with the instructor and adapted for students' self-assessment purposes as it was originally designed for instructors to use. As a result of modification, it's subject was changed from " The student" to "I", check boxes were put beside each descriptive line for the students to tick in upon achievement of the particular requirement and also some other modifications were made based on the task requirements. The writing task through which the training was given was argumentative writing. For that reason, an argumentative essay was used as a model paper by the researcher to demonstrate how the criteria will be applied to the assessment of the particular paper.

The instruments used for data collection were two questionnaires (background and attitude), all three drafts of the students' argumentative essays and the self-assessment criteria that the students were asked to use for each draft. Students were given a background questionnaire before the training and after the training they were given an attitude questionnaire to reflect their thoughts about the effectiveness of the training given and the practice of self-assessment. Training data were collected through each draft of the students' essays. The drafts were read and graded by two raters (their instructor and the researcher) and also self-assessment criteria was filled in by the students for each draft of the essay, on which they had also written their self-ratings for the three traits of the criteria and the total score.

For the analysis of the data collected through the study, primarily the background questionnaire results were examined to get initial information about the participants' age, gender, department, language level and writing experience.

Secondly, Pearson Product Moment Correlation formula was used to check the inter-rater- reliability between the two raters. Thirdly, Pearson Moment Correlation Formula and t-test were computed to see if the training caused the treatment group to perform significantly better in the self-assessment of their writing than the control group. At the next stage, comparison of the mean scores and t-test were used to find out whether the training caused improvement in the writing skills of the students in the treatment group. Finally, the attitude questionnaire, which was featured to elicit information on how the students in the treatment group felt about training offered on self-assessment of their writing skills, was again analysed by the use of statistical methods.

Results

In the data analysis, the inter-rater reliability between the scores of the external raters was found to be low, thus the scores of the two raters were combined to achieve higher reliability. In relation to the first research question, the data analysis showed that the ratings of the students in the treatment group increased constantly and consistently and got closer to that of the raters' ratings. Whereas the ratings of the students in the control group was inconsistent and different from that of the two raters. Still, the results were not statistically significant. The results of the data analysis for the second research question demonstrated that there was significant improvement in the writing skills of the students within both groups throughout the writing process. Still, there was no significant difference between the improvement in the writing skills of the students in the two groups as a result of the training. Finally, according to the results of the attitude questionnaire, the training was perceived to be effective by the majority of the students in the treatment group.

Students' answers to the attitude questionnaire revealed that their attitude towards self-assessment of their own writing is positive. Another important conclusion was that, the students were more comfortable using the self-assessment criteria than the generic criteria, and also assessing the model paper than assessing their own papers. Lastly, the data analysis of the attitude questionnaire revealed an important limitation of the study, that is, if the students were to receive the same training again, they would like to have more detailed instruction about scoring and longer time for training.

Discussion

As mentioned in the results section, the analysis of the data for the first question showed that the ratings of the students in the treatment group demonstrated a consistent increase and thus, got closer to that of the raters' ratings. Still, the results were not statistically significant as to say that the training caused important changes in the self-assessment results of the treatment group. There are several factors that may have affected the outcome of the training results. One factor may be the time constraints. Training the students on the grading criteria has been considered as a means to introduce self-assessment gradually (Ferris and Hedgcock, 1998; O'Malley and Pierce, 1996). Thus, if the training had been given over a longer period of time it may have been more effective, as it would have given the students the chance to internalize the criteria better. Another factor that affected the self-assessment results may be the type of self-assessment task. The students might have had difficulty self-assessing in an argumentative writing assignment whereas the results could have been different with another type of writing assignment. Lastly, small number of people in each group may be a possible reason why no significant result was

achieved. If the study is replicated with larger number of participants, it may yield more significant results.

The results of the data analysis also showed no significant improvement in the students' writing skills as a result of the self-assessment training. Again there may be several factors affecting the results. The students writing improvement could have been observed for each trait separately as suggested by Arter (1994), which may have yielded more useful results. The students may have performed better in one trait than the others. Another factor that made it difficult to observe the improvement in the students writing skills may be the process writing approach. In this approach, improvement in students writing skills is a natural consequence of the process. Also, continuous teacher feedback, which is an inherent part of process writing, may be another factor that made it more difficult to see the students' own improvement.

The analysis of the data collected through the attitude questionnaire, seem to confirm one important factor that may have affected the self-assessment results. The important reason for why they felt less comfortable scoring their own papers may be interpreted as the difficulty they had in being objective towards their own products.

Pedagogical Implications

There are several pedagogical implications that can be drawn from this study to be used in classroom instruction. Primarily, although not statistically significant, the findings showed changes in the self-assessment ratings of the students in the treatment group, in the direction of becoming closer to their instructors' assessment. The results can be seen as evidence that the training helped the students to develop a common understanding with their instructor on the set of criteria consisted in the

scoring rubric. Thus, engaging Freshman English students in self-assessment as part of the course may help the FYE instructors in their efforts to encourage learner autonomy and help their students' become aware of their strengths and weaknesses of their own product. Being involved in the assessment process and given training on the self-assessment of their writing skills, may also cause the students to develop a more realistic and objective attitude towards their own capacity and may eliminate the frustration caused by the teacher assessed final grade.

Similarly, even though in the present study no significant improvement was observed in the writing skills of the students, still as suggested in the literature and confirmed by the last questionnaire, the researcher believes training the learners on the self-assessment of their writing skills is a worthwhile endeavour that helps students to raise a critical awareness towards their own language abilities and language performance. Also, when it becomes part of the everyday classroom instruction, self-assessment may yield useful information both to the instructor and the students on their improvement within the course. This may also help to achieve better results, as the concern about the time constraints will have been eliminated.

Implications for Further Research

Results of this study suggest several considerations for further research. Researchers may think about repeating this study for a longer period of time and using simpler criteria, specified as much as possible. Also, another writing task may be used to see whether training on self-assessment of the students' writing skills would yield different results. Also, the students writing improvement can be observed separately for each trait in the assessment criteria used for training purposes, which may yield different and useful results. Repeating the research in

another writing approach other than process writing where nature of the writing method and teacher feedback would not interfere may be another consideration. Thus it may be easier to see the improvement in the students' writing skills.

REFERENCES

Arter, J. A. (1994). The impact of training students to be self-assessors of writing

[Online]. Available: ERIC Document Reproduction Service. ERIC Identifier:

ED370975.

Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford:

Oxford UP.

Brown, J. D. (1995). *The elements of language curriculum: A systematic approach to*

program development. Boston: Heinle & Heinle Publishers.

Brown, J. D., & Hudson, T. (1998). The alternatives in language assessment. *TESOL*

Quarterly,32, 657.

Brown, J. D. (Ed.). (1998). *New ways of classroom assessment.* Alexandria : TESOL

Publications.

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

Heinle & Heinle.

Dickinson, L. (1987). *Self-instruction in language learning.* Cambridge: Cambridge

UP.

Ekbatani, G., & Pierson, H. (Eds.). (2000).*Learner-directed assessment in ESL.*

Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Ferris, D., & Hedgcock, J. S. (1998). *Teaching ESL composition: Purpose, process*

and practice. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Grabe, W., & Kaplan, R. (1996). *Theory and practice of writing.* New York:

Addison Wesley Longman Ltd.

Hamp-Lyons, L. (Ed.). (1993). *Assessing second language writing in academic*

contexts. Norwood, NJ: Alex Publishing Corporation.

Harris, M. (1997). Self-assessment of language learning in formal settings. *ELT Journal, 51,* 12-20.

Hills, J. R. (1976). *Measurement and evaluation in the classroom.* Columbus, Ohio: Charles E. Merrill Publishing Company.

Hillocks, G. (1986). *Research on written composition. New directions for teaching.* Illinois: ERIC Clearing House Publications.

Hindman, J. E. (1994). Letting students in on the secrets of evaluation and placement [Online]. Available : ERIC Document Reproduction Service. ERIC Identifier: ED375402.

Huerta-Macias, A. (1995). Alternative assessment: responses to commonly asked questions. *TESOL Journal,5,* 8-11.

Janssen-van Dieten, A. (1989). The development of a test of Dutch as a second language: The validity of self-assessments by inexperienced subjects. *Language Testing, 6,* 30-46.

LeBlanc, R., & Painchaud, G. (1985). Self-assessment as a second language placement instrument. *TESOL Quarterly,19,* 673-687.

MacIntyre, P. D., Noels, K.A., & Clement, R.(1997). Biases in self-ratings of second language proficiency: the role of language anxiety. *Language Learning, 47,*265-287.

McNamara, J., & Deane, D. (1995). Self assessment activities: toward autonomy in language learning. *TESOL Journal,5,* 17-21.

Nunan, D. (1988). *The learner-centered curriculum.* Cambridge: Cambridge UP.

O'Malley, J. M., & Pierce, V. D. (1996). *Authentic assessment for English language learners.* Boston: Addison-Wesley Publishing Company, Inc.

Petrosky, A. R., & Bartholomae, D. (Eds.). (1986) . *The teaching of writing. Eighty-fifth yearbook of the national society for the study of education.* Chicago, Illinois: Chicago Press.

Pierce, B. N., Swain, M., & Hart, D. (1993). Self-assessment, French immersion and locus of control. *Applied Linguistics, 14,* 25-42.

Ross, S. (1998). Self-assessment in second language testing : a meta-analysis and analysis of experiential factors. *Language Testing, 15,*1-20.

Ross, J. A., Rolheiser, C., & Hogaboam-Gray. A. (1998). Effects of self-evaluation training on narrative writing [Online]. Available : ERIC Document Reproduction Service. ERIC Identifier: ED424248.

Smolen, L., Newman, C., Wathen, T., & Lee, D. (1995). Developing student self assessment strategies. *TESOL Journal, 5,* 22-27.

Stiggins, R. J. (1991). Facing the challenges of new era of educational assessment. *Applied Measurement in Education, 4,* 263-272.

Susser, B.(1994). Process approaches in ESL/EFL writing instruction. *Journal of Second Language Writing, 3,* 31-47.

Williams, J. D. (1998). *Preparing to teach writing: Reseach, theory and practice.(2nd ed.).* Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

APPENDIX A

Generic Writing Criteria
Engineering and Science
Fall 2000-2001

CONTENT 50%	
A 45-50	Student has full understanding and consideration of the rhetorical aspects of the task (Genre, Audience, Purpose, Message, Tone) Student has presented adequate, relevant and sufficient support for the ideas shown in the essay. Student has fully addressed the task and has adhered to the word limit assigned by the instructor. Student has consistently and effectively used the writing process by brainstorming a wide range of possible ideas, pre-writing ideas, drafting, and revising to develop and improve content. Instructor and peer review comments have been fully addressed.
B 39-44	Student has understood and considered most of the rhetorical aspects of the task (GAPMT) Student has presented partially adequate, relevant and sufficient support for the ideas shown in the essay. Student has not completely addressed the task and closely respected the word limit. In general student has effectively used the writing process to explore options and learn from them. However, the student may not have consistently taken advantage of the writing process. Instructor and peer review comments have been partially addressed.
C 33-38	Student has limited understanding and consideration of the rhetorical aspects of the task (GAPMT) Student has presented less than adequate, relevant and sufficient support for the ideas shown in the essay. Student has attempted to address the task but failed to cover all aspects. Student went over or under the word limit. There is limited evidence that student has effectively used the writing process. Brainstorming was limited to a few ideas, which were not explored in pre-writes or drafts. Limited revisions are seen in content. Instructor and peer review comments have been largely ignored.
D 28-32	Student has little or no understanding and consideration of the rhetorical aspects of the task (GAPMT) Student has presented poor or not adequate, relevant and sufficient support for the ideas shown in the essay. Student has made little or no attempt to address the task and failed to cover most of the aspects. Student did not adhere to the word limit. Student has simply written out drafts with no thought to exploring and learning through writing. Student showed no attempt to deal with ideas or expression. Little or no revision.
F 00-27	Completely irrelevant information. No indication of thoughtful writing.

ORGANIZATION 30%	
A 27-30	• Student has outstanding understanding of the genre conventions. The introduction effectively generates audience interest and establishes the writer's point of view. The conclusion clearly refers back to the introduction and summarizes essential points presented in the essay. • The relevance and importance of all ideas are clearly indicated. (use of transitions) • All paragraphs have a definite focus and transitions between paragraphs clearly inform the reader of the writer's intentions or point of view. • Student has consistently and effectively used the writing process to explore options and to present information in a creative and effective manner. Instructor and peer review comments have been fully addressed.
B 23-26	• Student has good understanding of the genre conventions. The introduction generates audience interest and establishes the writer's point of view. The conclusion somewhat refers back to the introduction and summarizes essential points presented in the essay. • The relevance and importance of most ideas are clearly indicated. (use of transitions) • Most paragraphs have a definite focus and transitions between paragraphs clearly inform the reader of the writer's intentions or point of view. • Student has in general effectively used the writing process to explore options and to present information in a creative and effective manner. Instructor and peer review comments have been partially addressed.
C 20-22	• Student does not have a clear understanding of the genre conventions. The introduction may generate little audience interest and may not clearly establish the writer's point of view. The conclusion does not refer back to the introduction and does not summarize essential points presented in the essay. • The relevance and importance of ideas are only partially indicated. (use of transitions) • Few paragraphs have a definite focus and transitions between paragraphs clearly inform the reader of the writer's intentions or point of view. • There is limited evidence that the student has effectively used the writing process. Limited revisions are seen in the organization of material; instructor and peer review comments have been largely ignored..
D 17-19	• Student does not have any understanding of genre conventions. The introduction lacks interest and is unfocused. The conclusion does not tie in with the essay at all. • The relevance and importance of ideas are not indicated at all. (use of transitions) • Paragraphs are completely disconnected. • Little or no revision is seen in organization.
F 00-16	• No apparent organization.

VOCABULARY/SENTENCE STRUCTURE/GRAMMAR/PUNCTUATION 20%	
A 18-20	Vocabulary, sentence structure, grammar, and punctuation are all mostly free from errors. Student has a wide range of word choice and diverse structures. Instructor and peer review comments have been fully addressed.
B 16-17	Vocabulary, sentence structure, grammar, and punctuation are in general free from errors and no errors prevent communication. Student has made noticeable effort to acquire a wide range of word choice and diverse structures, but does not have full command of the new material. Instructor and peer review comments have been partially addressed.
C 14-15	There are several vocabulary, sentence structure, grammar, and punctuation errors some of which interfere with communication. Student has made some effort to acquire a wide range of word choice and diverse structures, but has only limited command of the new material. Instructor and peer review comments have been hardly addressed.
D 12-13	Many vocabulary, sentence structure, grammar, and punctuation errors prevent communication. Student has made little or no effort to acquire a wide range of word choice and diverse structures, but does not any command of the new material. Instructor and peer review comments have not been addressed at all.
F 00-11	Incomprehensible.

APPENDIX B

There is Someone Inside Me

There was a mysterious atmosphere in the room. The gloomy lights were making the mystery more and more attractive but terrifying also. And at that time, two strong lights appeared suddenly: One of them was blue and the other was red. Then the lights collided and the bodies came into view. Blue light belonged to the white cloth and the red light belonged to the black cloth. These bodies started to fight with their swords, light sabres, and an incredible voice covered everywhere. Yes, Luke and Darth Vader were fighting; there was nobody around them except the mysterious atmosphere. They were going one step back and forward. When you started to think that the battle would never end, Luke made a clever attack and Darth Vader lost his light sabre so he was completely defenseless from Luke. Luke didn't miss this chance and killed him. Darth Vader fell and gave his last breath. Luke managed to overcome his enemy, at last. He was in a big curiosity about Darth Vader. Then he decided to learn the face under the mask. He started to lift the mask slowly and took it at last. Oh my God!...What was it? He didn't beieve his eyes because the face wasn't anyone else' face but Luke!... He woke up then and thought himself. What did it mean? (Star Wars Episode 2)

The small part of the film tells a lot of things to us in fact. Luke came acrss such a situation which left him in a big astonishment. His enemy was near him as himself. He was carrying the dark side in his body. Or you can think the other way: Is Luke the good side of Darth Vader? I think there is no need to think that it is too surprising because every good has a bad side and every bad has a good side in him or

her. From the beginning of humanity, there has been a battle between good and evil. This was the nature's rule and it will be. While life continues like this, people join a group when they reach an age. They do this operation by their concience, society and society's expectations. Consider heroes and villains. Being a villain is of course much profitable. Villains have some real incentives to be villains: money, love slaves, armies, global domination, and the freedom to be their own boss. Heroes, on the other hand, have little incentive: respect, admiration, limited freedom. So the villians have a lot of comfort. They can rape, ruin, steal, break laws, take lives. On the other hand, heroes are servants to society and they must be benovelent, generous and honest. But a person can't be completely evil or completely good. Although he or she is good, nobody can be excellent. Perfection and being a human are two different things. And also being a human-being brings the emotions. A weak point always exists and the worst person has a love to a human or being certainly. This situation is strongly seen in Chinese philosophy under the name of yin and yang. Now let's have a look at this concept.

The twin concepts of yin and yang are symbolised by the sun and the moon. They are two opposing forces active in the universe. Yin exists in yang an yang exists in yin. This changing combination of negative and positive, dark and light, cold and hot give the force of the universe. Yin symbolizes female , earth, moon, dark, femine, negative, absorbing, receptive, winter, soft, cold, even and so on. And yang symbolizes male, heaven, sun, bright, fiery, moving, odd, powerful, active, summer and so on. Heaven was created by yang and the earth was created by yin, yang provides peace and yin makes confusion. Water is a product of yang and yang builds the air, as yin builds the senses. Sense belongs to the body. When body dies,

the sipirit remains in the air and takes its' food from the air and the body receives its'

food from the senses. If yin is very strong, yang can be too weak. And if yang is very

strong, yin can remain weak. The symbol of yin and yang (Thai-Chi) shows each

element has a small piece of the other within it. When one of them reaches its'

extreme point, the other one's influences decrease but there is a balance between

them and these concepts never disappear. This situation resembles the dominant and

recessive genes. Both of them exist but they don't show themselves every time. And

also this can remind us that a person rarely notices all of the features but a person can

notice the most visible features.(http://www.paei.com/DN/WU/WU94/Modo.html)

The eleventh century (AD) philosoph Chu His explains this condition as:

> The yang transforms and the yin conserves. The yang and the yin manifest as
>
> movement and rest; yang moves to its utmost, then rests; yin rests to its
>
> utmost, then moves. Therefore, yin rests within yang moves within yin; the
>
> two are inseparably interwoven. It is thus a single unit that they are one with
>
> the Tao. (http://www. macellan.com)

I think this quotation explains the positions of yin and yang. He claims that in

spite of being opposite, yin and yang complete eachother. In my opinion Chu his has

pointed out the main topic precisely. Because we can't separate mixed things and yin

and yang carry some parts from eachother.

The reality takes its placein Chinese philosophy in this way. And we also see

the same pattern in fantastic literature. Fantastic literature always contains the good

and the evil wars. As talks about this opposite concepts, it refers to the weak points

of goods. J.R.R Tolkien created a character named Thorin in Hobbit. Thorin is a

dwarf but he is keen on money very much. This is his weak point. When Bard and the Lake-men want the recompense of their damage, Thorin replies as follows:

> You put your worst cause last and in the chief place. To the treasure of my people no man has a claim, because Smaug who stole it from us also robbed him of life or home. The treasure was not that his evil deeds should be amended with a share of it. The price of the goods and the assistance that we received of the Lake-men we will fairly pay-in due time. But nothing will we give, not even a loaf's worth, under threat of force. While an armed host lies before our doors, we look on you as foes and thieves. It is in my mind to ask what share of their inheritance you would have paid to our kindred, had you found the hoard unguarded and slain. (Tolkien 237)

By these words, Thorin wants to say that the treasure belonged to Thorin's ancestors and now, it belongs to Thorin. So he thinks he doesn't have to give a part of it to Lake-men. He doesn't feel like he has to give a recompense to them.

The wish which the lake man wanted was right, I think. Because Smaug gave a very bad damage to the Lake-men's town. A lot of people lost their lives also. And Smaug was disturbing them before in any case. So Thorin had to pay the harm but he didn't want to give any money to them. He behaved like a selfish person. That is, in spite of being a good guy, he has evil points in his character.

Because of the concepts good and evil are common in social ilfe, social concepts, certainly, include these topics. Especially religion stress them. For example; Christians believe that all people are guilty by birth. Because Adam and Eve shared a sin together by eating the forbidden apple and then they became guilty. Because each person is their son or daughter, all people have the same sin by birth.

Also our religion, Islam, has a related concept with good and evil. Each person is created according to the Islam Creation but the servant (to the God) has desire also. This desire causes the servant to commit sins. And because of being created according to the Islam Creation, an evil person can't be completely bad. He or she has a tendency to be a good person. So there are a lot of ways which can providethe transition between the good and the bad. The choice depends on the servant.

As we can see, the concepts good and bad have a great influence on themselves. They carry some parts from eachother. This is valid for every person. The only thing which can be changeable from person to person is the parts' size, I think. Everything around us can demonstrate this to us. This is a fact and stays as a fact forever.

APPENDIX C

Informed Consent Form

Dear Student,

You are being asked to give me permission to use your essay as a model paper in an experimental research study. The aim of the study is to explore the effect of training the First Year English students at Bilkent University on the self-assessment of their own writing. To be able to give the training, I need to use a model paper on which the scoring process will be demonstrated.

Your permission of the use of your paper as a model essay will highly contribute to my study. Any information given to me as the researcher will be kept confidential and your name will not be released. I would like to thank you for your contribution in advance. If you have any questions please do not hesitate to contact me at the phone number given below.

Very Truly Yours,

İKLİL KAYA YILDIRIM

MATEFL Program
Bilkent University
Ankara
Phone: 312-290 14 26

I have read and understood the information given above. I hereby agree to let the researcher use my essay for training purposes in her research.

Name:

Department:

ID Number:

Signature:

Date:

APPENDIX D

Out-of-Class Essay #2: Analysis and Response to a Text (25%)

For your second essay you will respond to ideas presented in *The Hobbit* by J. R. R. Tolkien. This will be a more formal academic essay than the first. In it you will accurately summarize, and intelligently respond to idea(s) from that text. In this response, you will make a claim and then support it with your own ideas as well as the ideas of others. You will refer to any course text through paraphrase and/or direct quote *at least* three times in your essay, and don't forget to follow each reference with the page number where the quote or paraphrased text can be found. Your references will demonstrate that you are able to apply ideas learned previously in this class to the reading of a text.

Why should I write an essay like this? What will I get out of it?

Basically, academic writing almost always involves some kind of summary, synthesis, analysis, and response/reaction/application/argument. We construct and develop our viewpoints based on written and/or spoken texts. When we build arguments about 17th century novels, the movements of quarks in atoms, and the state of SE Asian economies, we demonstrate our credibility by showing that we have read what has been written on the topic, understood it, and are capable of criticizing and extending it. These texts provide the foundation upon which we begin our own research.

One possible way to organize this essay:

- In the introduction of this essay, you will summarize relevant aspects of the source text, in your case *The Hobbit*. Your summary will show that you have clearly understood and *analyzed* the author's intention and how

this is portrayed in the text. You should then make clear how this leads to your own claim. This is important whether you plan to criticize or agree with the ideas--your summary must show that you understand the ideas before you go on to deny or affirm their validity. As an introduction, the summary should be less than one third of your essay.

- Your summary will lead in to your own claim about the ideas in the text in the form of an *explicit thesis statement*. This claim gives you something to prove in the body of your essay, as well as beginning to set up the organization of your essay.

- In the second part of this essay, you will defend the claim(s) you have made in your thesis statement. You may do this by citing evidence from your personal experience, but mainly you will apply structures and ideas you have learned throughout the course to the ideas you have summarized from the primary text.

- In the conclusion of your essay, you have the chance to demonstrate the sophistication of your view. You have shown your clear understanding of the text in your introduction, you have constructed an intelligent response in the body of your essay; now it is time to answer the question: "So what?" Look to the future-- why is your argument important? Where should we go from here?

Academic skills developed in this essay:

Revision of skills from essay #1 (focus, detail, specific evidence, avoidance of general language, etc.)

⇒ Summarizing to demonstrate understanding of an author's intention through use of academic vocabulary (The writer *demonstrates, proposes, describes, clarifies, illustrates...*).

⇒ Smooth integration of source material into a text through paraphrase and direct quote.

⇒ Developing a coherent and persuasive reaction to a written text that reflects/applies previous course content.

Technical details:

- Hand in two copies of the final essay.

- Please include all previous drafts and prewriting material with each paper turned in.

- Every draft of your paper will be 4 pages long, typed on a computer in 11 or 12 pt. font (font must be easy to read—we suggest Ariel or Times New Roman), double spaced, and with appropriate margins (usually 2.5 to 3 cm on all sides)— it has got to look good.

- First draft due week 10 (December 7, 2000)

- Second Draft due week 12 (exact date to be announced)

- Final draft due week 15 (exact date to be announced)

Adapted from PD (Professional Development) Person, First-Year English Program, Bilkent University.

APPENDIX E

SELF-ASSESSMENT CRITERIA

CONTENT 50%
A   45-50

❏I have full understanding and consideration of the rhetorical aspects of the task (Genre, Audience, Purpose, Message, Tone)

❏I have presented adequate and relevant support for the ideas I show in my essay (at least 3 references/examples)

❏I have fulfilled all the requirements asked in the task and have adhered to the word limit (at least 4 pages) assigned by the instructor

❏I have consistently and effectively used the writing process by brainstorming a wide range of possible ideas, pre-writing ideas, drafting, and revising to develop and improve content. I have fully addressed instructor and peer comments.

B 39-44

❏I have understood and considered most of the rhetorical aspects of the task (GAPMT)

❏I have presented partially adequate and relevant support for the ideas I show in my essay. (at least 3 ref/eg)

❏I have not completely addressed the task and adhered to the word limit.(at least 4 pages)

❏In general I have effectively used the writing process to explore options and learn from them. However I may not have consistently taken advantage of the writing process. I have partially addressed instructor and peer review comments.

C 33-38

❏ I have limited understanding and consideration of the rhetorical aspects of the task (GAPMT).

❏ I have presented less than adequate and relevant support for the ideas I show in my essay.(at least 3 ref/eg)

❏ I have attempted to address the task but failed to cover all the aspects. I went over or under the word limit.(at least 4 pages)

❏There is limited evidence that I've effectively used the writing process to explore options and to learn from them. My brainstorming was limited to a few ideas, which I did not explore in pre-writes or drafts. I have made limited revisions in content. I have largely ignored instructor and peer review comments.

D 28-32

❑ I have little or no understanding of the rhetorical aspects of the task (GAPMT).

❑ I have presented poor or none adequate and relevant support for the ideas i show in my essay. (at least 3ref/eg)

❑I have made little or no attempt to address the task and failed to cover most of the aspects. I did not adhere to the word limit (at least 4 pages).

❑ I have simply written out drafts with no thought to exploring and learning through writing. I showed no attempt to deal with ideas or expression. I have made little or no revision.

F 00-27

❑ Completely irrelevant information. No indication of thoughtful writing.

ORGANIZATION 30%

A 27-30

❑I have outstanding understanding of the genre conventions. My introduction effectively generates audience interest and establishes my point of view. My conclusion clearly refers back to the introduction and summarizes essential points I have presented in the essay.

❑I have clearly indicated the relevance and importance of all ideas (use of transitions)

❑All of my paragraphs have a definite focus and transitions between paragraphs clearly inform the reader of my intentions or point of view.

❑ I have consistently and effectively used the writing process to explore options and to present information in a creative and effective manner. I have fully addressed instructor and peer review comments.

B 23-26

❑I have good understanding of the genre conventions. My introduction generates audience interest and establishes my point of view. My conclusion somewhat refers back to the introduction and summarizes essential points I have presented in the essay.

❑I have clearly indicated the relevance and importance of most ideas (use of transitions).

❑Most of my paragraphs have a definite focus and transitions between the paragraphs clearly

inform the reader of my intentions or point of view.

❑I have in general effectively used the reading process to explore options and to present information in a creative and effective manner. I have partially addressed instructor and peer review comments.

C 20-22

❑I do not have a clear understanding of the genre conventions. My introduction may generate little audience interest and may not clearly establish my point of view. My conclusion does not refer back to the introduction and does not summarize essential points I have presented in the essay.

❑ I have only partially indicated the relevance and importance of ideas. (use of transitions)

❑Few of my paragraphs have a definite focus and transitions between paragraphs clearly inform the reader of my intentions or point of view.

❑There is limited evidence that I have effectively used the writing process. I have made limited revisions in the organization of the material. I have largely ignored instructor and peer review comments.

D 17-19

❑I do not have any understanding of genre conventions. My introduction lacks interest and is not focused. My conclusion does not tie in with the essay.

❑ I have not indicated the relevance and importance of ideas at all (use of transitions).

❑My paragraphs are completely disconnected.

❑I have made little or no revision in the organization.

F 00-16

❑ No apparent organization.

VOCABULARY/SENTENCE STRUCTURE/PUNCTUATION 20%

A 18-20

❑My vocabulary, sentence structure, grammar and punctuation are all mostly free from errors.

❑I have a wide range of word choice and diverse structures.

❑I have fully addressed instructor and peer review comments.

B 16-17

❑My vocabulary, sentence structure, grammar, and punctuation are in general free from errors and no errors prevent communication.

❑I have made noticeable effort to acquire a wide range of word choice and diverse structures, but I do not have a full command of the new material.

❑I have partially addressed instructor and peer review comments.

C 14-15

❑I have several vocabulary, sentence structure, grammar and punctuation errors some of which interfere with communication.

❑I have made some effort to acquire a wide range of word choice and diverse structures, but I have only limited command of the new material.

❑I have hardly addressed instructor and peer review comments.

D 12-13

❑I have many vocabulary, sentence structure, grammar and punctuation errors that prevent communication.

❑I have made little or no effort to acquire a wide range of word choice and diverse structures, but I do not have command of the new material.

❑I have not addressed instructor and peer review comments at all.

F 00-11

❑Incomprehensible.

CONTENT SCORE   --------------------

ORGANIZATION SCORE   --------------------

MECHANICS SCORE   --------------------

TOTAL SCORE   --------------------

APPENDIX F

Out-of-Class Essay #2: Analysis and Response to a Text (25%)

For your second essay you will respond to ideas presented in *Sons and Lovers* by D. H. Lawrence. This will be a more formal academic essay than the first. In it you will accurately summarize, and intelligently respond to idea(s) from that text. In this response, you will make a claim and then support it with your own ideas as well as the ideas of others. You will refer to any course text or text(s) from the library or the internet through paraphrase and/or direct quote *at least* three times in your essay, and don't forget to follow each reference with the page number where the quote or paraphrased text can be found. Your references will demonstrate that you are able to apply ideas learned previously in this class to the reading of a text.

Why should I write an essay like this? What will I get out of it?

Basically, academic writing almost always involves some kind of summary, synthesis, analysis, and response/reaction/application/argument. We construct and develop our viewpoints based on written and/or spoken texts. When we build arguments about 17th century novels, the movements of quarks in atoms, and the state of SE Asian economies, we demonstrate our credibility by showing that we have read what has been written on the topic, understood it, and are capable of criticizing and extending it. These texts provide the foundation upon which we begin our own research.

One possible way to organize this essay:

- In the introduction of this essay, you will summarize relevant aspects of the source text, in your case *Sons and Lovers*. Your summary will show that you have clearly understood and *analyzed* the author's intention and

how this is portrayed in the text. You should then make clear how this leads to your own claim. This is important whether you plan to criticize or agree with the ideas--your summary must show that you understand the ideas before you go on to deny or affirm their validity. As an introduction, the summary should be less than one third of your essay.

- Your summary will lead in to your own claim about the ideas in the text in the form of an *explicit thesis statement*. This claim gives you something to prove in the body of your essay, as well as beginning to set up the organization of your essay.

- In the second part of this essay, you will defend the claim(s) you have made in your thesis statement. You may do this by citing evidence from your personal experience, but mainly you will apply structures and ideas you have learned throughout the course to the ideas you have summarized from the primary text.

- In the conclusion of your essay, you have the chance to demonstrate the sophistication of your view. You have shown your clear understanding of the text in your introduction, you have constructed an intelligent response in the body of your essay; now it is time to answer the question: "So what?" Look to the future-- why is your argument important? Where should we go from here?

Academic skills developed in this essay:

⇒ Revision of skills from essay #1 (focus, detail, specific evidence, avoidance of general language, etc.)

⇒ Summarizing to demonstrate understanding of an author's intention through use of academic vocabulary (The writer *demonstrates, proposes, describes, clarifies, illustrates...*).

⇒ Smooth integration of source material into a text through paraphrase and direct quote.

⇒ Developing a coherent and persuasive reaction to a written text that reflects/applies previous course content.

Technical details:

- Hand in two copies of each draft.

- Please include all previous drafts and prewriting material with each paper turned in.

- Every draft of your paper will be 4 pages long, typed on a computer in 11 or 12 pt. font (font must be easy to read—we suggest Ariel or Times New Roman), double spaced, and with appropriate margins (usually 2.5 to 3 cm on all sides)— it has got to look good.

- First draft due week 11 (April 16, 2001)

- Second Draft due week 13 (exact date to be announced)

- Final draft due week 15 (exact date to be announced)

Adopted from PD (Professional Development) Person, First-Year English Program, Bilkent University.

APPENDIX G

Dear Student,

This questionnaire was designed for the research I'm conducting at Bilkent

University Freshman Department. The results of the questionnaire will be used only

in this research and kept confidential. Thank you for your cooperation.

AGE:------------

GENDER : MALE -------FEMALE -------

Department : --------------------------

1. Have you studied English in BUSEL ?

   ❑ Yes   ❑ No

2. If your answer to the first question is Yes, then could you specify how long

   you've studied in BUSEL? Please write the number of semesters in the given

   space.

   ------------------ semester/s

3. If your answer to the first question is No, then could you please circle the choice

   that best represents your present situation.

   ❑I'm a direct entry.

   ❑I'm repeating the course.

   ❑I'm on probation.

4. Have you had any writing courses before?

   ❑Yes   ❑ No

5. Have you ever practiced process writing (writing essays in several drafts) before?

   ❑Yes   ❑ No

6. Was the writing experience you had in BUSEL similar to what you have in the Freshman Department now?

❑Yes   ❑ No

7. Can you identify the strengths and weaknesses of your paper and decide on your writing quality?

❑Yes, I can.

❑No, I can't.

❑I can realize them only after the teacher's/peer's feedback.

8. Could you estimate the grade you'll get from your writing assignments?

❑always   ❑seldom

❑usually   ❑never

9. If you could estimate your grade, then are you

❑just guessing?

❑checking your essay against the writing prompt?

❑checking your essay against the grading criteria?

❑deciding according to the teacher's feedback?

10. Do you take your essays to Bilwrite Center for getting feedback other than your teacher's or your peer's feedback?

❑always   ❑seldom

❑usually   ❑never

APPENDIX H

Dear Student,

This questionnaire has been designed for a research study conducted at Bilkent University First Year English Program. The results of this questionnaire will be used to evaluate the effect of the training offered on your self-assessment of your own writing skills. Thank you for your cooperation.

1. Do you think the training given to you on the self-assessment of your papers was effective?

   a) Yes   b) No

2. If your answer to the first question is Yes, then please circle the most appropriate reasons for why you think the training was effective? Check all that apply.

   f) I developed the habit of using the criteria effectively.

   g) I can now see the strengths and weaknesses of my paper.

   h) I can now predict the points on which my teacher will give feedback.

   i) I can now make realistic guesses about the grade I will receive.

   j) I believe that I learned how to self-assess.

3. If your answer to the first question is No, then please circle the most appropriate reasons for you out of the below given choices? Check all that apply.

   f) I still can't use the criteria effectively.

   g) I still can't see the problems in my paper.

   h) I can't trust in my own assessment.

   i) I can't guess the points that I will get feedback on from my teacher.

   j) I don't think that I learned how to self-assess.

4. Would you like to go on with self-assessment by yourself in your future writing assessments?

   a) Yes   b) No

5. What do you think was the most useful part of the training on self-assessment of your own writing? Please circle the answers that best express your thoughts. Check all that apply.

   a) learning how to use the criteria

   b) having control over the assessment of my paper

   c) learning how to self-assess

   d) learning how to identify the problems in my paper and to revise them

   e) improving my writing

6. Which essay was more difficult to assess for you?

   c) the model essay used in the first session of the training.

   d) my own essay.

7. Which criteria were you more comfortable using?

   c) generic writing criteria the subject of which was " The student "

   d) self-assessment criteria the subject of which was " I "

8. What do you think were the positive (+) and the negative (-) aspects of the self-assessment criteria? Put a check in the box that best reflects your thoughts about each item on the left.

	Positive	Negative
a) the subject " I "	☐	☐
b) the check boxes beside each descriptive line	☐	☐
c) the categories of the criteria (content, organization,	☐	☐

vocabulary/sentence structure/mechanics)

d) the language of the criteria   ☐   ☐

e) the practicality of scoring   ☐   ☐

9. Which of the three traits (content, organization, mechanics) was more difficult to score?

    d)   content

    e)   organization

    f)   mechanics

10. If you were to receive another training on the self-assessment of your writing

     skills, which of the following would you like to have ?

    d)   simpler criteria

    e)   longer time for training

    f)   more detailed instruction about scoring

APPENDIX I

First-Year English Program Principles & Suggestions

Critical Thinking Goals:

to engage students in their own critical, creative and reflective learning process by:

- allowing students to actively contribute to the course – both in terms of content, assignments, assessment, and feedback – in pedagogically sound ways;

- encouraging students to monitor their own learning processes;

- promoting learner autonomy by broadening the focus of courses to include exploration and discovery as fundamental curricular goals;

- designing courses and materials that encourage the consideration of perspectives different from those commonly held by students;

- encouraging students to make connections within and outside of class in creative, critical ways;

- designing materials that encourage students to articulate for themselves and others their own convictions and to reflect on how their thought changes throughout the course;

- creating a learning environment which validates students' ideas and encourages them to respect and engage each other;

- providing one-to-one and small group tutorials.

Principles:

In Bilkent First Year English 101 and 102, instructors will:

- administer at least one mid and one end of semester written course evaluation

that is specific to the course.

- provide student feedback on course content that will be given serious consideration in course revision.

- join their unit in compiling, analyzing and sharing feedback and use it in revising/redesigning courses.

<u>Suggestions:</u>

Bilkent First-Year English 101 and 102 students will engage in their own critical, creative and reflective learning process (below are some suggestions):

❏ contributing to the course – both in terms of content, assignments, assessment, and feedback – in pedagogically sound ways by addressing some number of the following.

<u>Content and feedback:</u>

- choosing texts from a group selected by the teacher.

- taking part in courses that lead gradually to content which content is entirely selected by the students.

- taking part in student 'jam sessions' (e.g., informal brainstorm-like sessions) to collect more informal oral feedback, including using a class representative to present students ideas, etc.

<u>Assignments:</u>

- contributing to refining/defining assignments provided by the teacher.

- choosing between several options of assignments.

- creating/developing an assignment as a part of the assignment.

  <u>Assessment:</u>

- analyzing and/or evaluating a given assignment criteria.

- creating a criteria for assessment.

- designing quiz, reading test, or other assessment tools.

❑ monitoring their own learning processes.

- engaging in self-assessment as part of the course.

- producing reflective writing tasks that ask them to consider particular processes of learning they have experienced and how they have developed through them.

- writing reflective diaries/journals.

❑ connecting classroom experience with social and personal experience in creative, critical ways.

- inviting speakers to the class.

- bringing student generated materials to class

- doing interviews and polling

- making use of Internet resources, such as chat rooms, listserves and email pen-pals.

- using role-play

- relating discussion topics to the Turkish experience

- attending field trips

❑ experiencing learner autonomy by taking courses that broaden the focus of learning to include exploration and discovery as fundamental curricular goals.

- taking part in student-initiated independent projects such as finding out about opera, for example, in Ankara and writing a report about it.

- researching/exploring topics around some type of independent study.

completing assignments that emphasize discovery and uncertainty in addition to traditional argument and problem/solution types of tasks.

Bilkent First-Year English 101 and 102 instructors will engage students in their own critical, creative and reflective learning process by addressing some number of the following:

- ❑ designing courses and materials that encourage the consideration of perspectives different from those commonly held by students.

  - creating courses whose subject matter covers a variety of cultural backgrounds and perspectives.

  - designing text-based assignments that ask students to examine how and why their own beliefs differ from others'.

  - selecting content areas of local and global importance that students may not be very familiar with and that will challenge and complicate beliefs and values they may have developed throughout their lives.

  - building into the course texts and activities that address the more general notion of "change" and human development (e.g., what is gained and lost by value shifts? Why are both tradition and change valuable to an individual and society?).

  - adopting teaching practices that are intellectually challenging, but not emotionally threatening. For example, a sensitive role play activity ("What would you do if your best friend told you he/she was gay?") needs to be done in a tolerant atmosphere in which students know they will be listened to.

- ❑ creating a learning environment which values students ideas and encourages them to respect and engage each other.

- discussing classroom management and operation with the students

- addressing teacher/student expectations explicitly

- modeling and encouraging active listening skills

- teaching functionally appropriate language forms (e.g., agreeing, disagreeing, clarifying, etc.).

- facilitating activities such as debate, discussion and presentation.

❑ designing courses and materials that encourage students to articulate for themselves and others their own convictions and to reflect on how their thought changes throughout the course.

- assigning and using for reflection pre- and post-unit writing tasks (or video mini-presentations, or audio recordings) where students express their positions on course content issues, and how their beliefs have been influenced by encounters with those who think differently from them.

- having students keep Journals to reflect on how their thinking has changed throughout the duration of the term. Periodically, giving reflective assignments asking students to do the same.

- perhaps, earlier in the term, using texts and activities which examine the value of self-reflection and how certain authors use writing to reflect on their lives (again, not to limit the medium, since they could take forms other than writing).

- using texts featuring characters who have undergone fairly dramatic changes (e.g., *The Autobiography of Malcolm X*, *Metamorphosis*, etc.).