Effectiveness of Self- Assessment Technique in Reading Skills of EFL Students

Mashael Nasser A Bin Rabiah

Department of English Language, College of Applied Studies and Community Service, King Saud

University

mrabeah@ksu.ed.sa

ABSTRACT

This study investigated the effectiveness of the Self-Assessment Technique (SAT) in developing the English reading skills of female students at the College of Community Service and Applied Studies in Riyadh. The participants of the study were students in first-year classes. They were equally divided into two groups: an experimental and a control group with 20 participants in each group. Teaching the experimental group involved using the SAT, while the control group was exposed to a conventional method of assessment. The study lasted a period of eleven weeks with eight weeks allocated for the intervention. Prior to and at the end of this period, students of both groups were tested on reading skills. The study found that the two groups had equal mean scores on the reading skills pre-test. The experimental group had a higher mean score on the reading skills post-test, but the difference was not significant at the 0.05 level. Based on these results, conclusions were drawn and recommendations were made.
Keywords: English language learning, language skills, self-assessment, reading strategies, reading comprehension.

ملخص البحث

هدفت هذه الدراسة إلى التحقق من مدى فعالية استخدام تقنية التقييم الذاتي في تعلُّم مهارات القراءة باللغة الإنجليزية لطالبات كلية الدراسات التطبيقية وخدمة المجتمع بالرياض. شارك في هذه الدراسة طالبات السنة الأولى حيث تم تقسيم عينة الدراسة بالتساوي إلى مجموعتين هما مجموعة تجريبية ومجموعة ضابطة مع ٢٠ مشاركة في كل مجموعة. درست المجموعة التجريبية مهارات القراءة باستخدام طريقة التقييم الذاتي بينما المجموعة الضابطة درست المهارات نفسها بالطريقة التقليدية للتقييم. استغرقت هذه الدراسة فترة زمنية مدتها أحد عشر إسبوعاً خُصص منها ثمانية أسابيع للتعلم بطريقة التقييم الذاتي. وقد أُجري اختبارين في مهارات القراءة لطالبات المجموعتين على حدٍ سواء إحداهما قبلي و الآخر بعدي. كشفت نتائج الدراسة أن كلتا المجموعتين حصلتا على متوسط درجات متساوٍ في اختبار مهارات القراءة القبلي ،كما أظهرت أن المجموعة التجريبية التي درست بطريقة التقييم الذاتي حققت متوسط درجات أعلى في الاختبار البعدي لمهارات القراءة إلّا أن هذا الفرق لم يكن ذا دلالة إحصائية عند مستوى الدلالة ٠,٠٥ .بناءً على هذه النتائج تم استخلاص الاستنتاجات و التوصيات.
الكلمات المفتاحية: تعليم اللغة الانجليزية، مهارات اللغة، التقييم الذاتي، استراتيجيات القراءة، استيعاب المقروء.

Introduction

An important aspect of teaching is assessment or evaluation. According to Harris and Mccann, (1994) in assessment, we measure student performance, look at their progress, diagnose their problems, and provide them with feedback. One technique that has surfaced in recent times is SAT. M. Oscarsson (1997) defined Self-Assessment as a technique that refers to how, under what conditions, and with what effects learners and users of a foreign or second language may judge their own ability in the language. In language teaching, Self-Assessment can be a beneficial technique because it provides teachers with a vital source of information about students' needs, expectations, and problems. Researchers (Marhaeni, 2005; Harris, 1997; M. Oscarson, 1989) and teachers around the world realize its potential benefits for education and have begun to use it and find ways of incorporating the Self-Assessment Technique in general into the educational process especially in the teaching of a second or foreign language such as English. However, the Self-Assessment approach has not yet come into actual use in KSA schools and universities. Thus, this study is carried out to investigate its potential effectiveness in language learning and teaching in KSA.

Statement of the Problem

In Saudi Arabia, English is used as a foreign language for international business, communication purposes and it is seen as essential for different jobs. Moreover, the ability to read and comprehend the English language effectively is considered critical to successful performance in numerous occupations, including language translation. In the English Translation Department at the College of Community Service and Applied Studies of King Saud University, the successful performance of such tasks as translating information of different topics and understanding spoken or written discourse depends to a considerable degree on language proficiency and skills in reading. Yet, students here and in Saudi universities, in general, do not show the required English proficiency according to Alsamadani (2009), and this could be due to their poor English especially reading skills and ineffective reading assessment.

From my own experience, I find that many teachers, who have taught the reading courses to students majoring in translation at the College of Community Service and Applied Studies at King Saud University, often complain about the students' poor performance. In this college, more often reading skill is taught using the lecture approach and assessment generally centers on the traditional method of assessment such as tests and assignments. Also, at the end of the semester, the students' English comprehension usually does not show satisfactory outcomes judging from the rather high

and consistent percentage of failures and the low scores of monthly and final examinations. This would be an indicator that the students are not internalizing what they have learned in the normal manner of teacher and assessment. Because of the major role of reading skills at the Translation Department of KSU's Community College, it is important to diagnose student problems in acquiring and developing those skills as well as improving their general language performance. By looking at methods of assessment and moving on to the use of the Self-Assessment approach, this would make the students become the focus of the learning process. In the Self-Assessment Technique, the students are able to evaluate themselves and to identify and become more aware of their problems. This will give teachers clues about the certain difficulties the learners are facing and provide directions on how to help them perform better.

Purpose of the Study

The purpose of the study is to investigate the effectiveness of Self-Assessment Technique of the English reading skills among students, and for this study, the focus is on the female students at the College of Community Service and Applied Studies in Riyadh. It particularly seeks to determine the effect of Self-Assessment tools such as the One-Minute Paper (Appendix A) and the Rating-Scale Sheet (Appendix B) on student reading comprehension and strategies.

Research Questions

This study was guided by the following questions:

1. How does the Self-Assessment Technique affect the reading skills (comprehension and strategies) of the experimental group compared to the control group?

2. Is there any significant improvement in the reading skills of the two groups?

Research Null Hypotheses

The following list shows the hypotheses of the study:

1. The Self-Assessment Technique does not affect the performance of the experimental group in reading skills (comprehension and strategies).

2. There is no significant improvement in the reading skills of the two groups

Definitions of Related Terms

Self -Assessment Technique (SAT)

The study adopted Marhaeni (2005)'s definition of Self-Assessment Technique in the English as a foreign language context that allows students to develop their ability to think for themselves and to evaluate what they learn. The present study is confined to two Self-Assessment tools which are the One- Minute Paper (Appendix A) and the Rating- Scale Sheet (Appendix B).

 Reading Comprehension

The study adopted Klingner, Vauughn and Boardman's definition (2007) of reading comprehension as "the process of constructing meaning by coordinating a number of complex processes that include word reading, word and world knowledge and fluency. It is a multi component, highly complex process that involves many interactions between readers and what they bring to the text "(p. 7). For the present study, this term refers to the reading comprehension skills including word meaning, word decoding, phonological awareness, vocabulary, and grammatical knowledge.

 Reading Strategies

For the present study, this term refers to the classification of reading strategies into three types: pre-reading strategies, while reading strategies and post-reading strategies. The study also adopted Alsamadani (2009)'s definitions of the three reading strategies.

Pre-reading strategies: are strategies used by readers before reading texts such as setting a purpose for reading and guessing what the text will be about.

While reading strategies: are strategies used by readers while reading texts such as keeping concentration focused, making inferences, and engaging with the text.

Post-reading strategies: are strategies used by readers after reading texts such as judging the consistency of the new information, modifying and evaluating new learning, and linking new learning with existing knowledge.

Literature Review

The Importance of SAT in English Language Learning and Teaching

Different studies of Self-Assessment have shown the positive aspects of this strategy and how it contributes to teaching and learning. Marhaeni (2005) explained that when students evaluate their performance, the result encourages them to set higher goals and commit more effort.

This combination of goal and effort equals achievement, then, a student's achievement results from self-judgment. Therefore, all can combine to impact self-confidence in a positive way. Through Self-Assessment, learners can monitor their own progress in terms of learning a foreign language. By following their own progress, Harris (1997) emphasized that they will eventually be able to realize that studying languages is different from other kinds of learning at their school or university. In learning a foreign language, the main objective is performance in the language rather than knowledge about the language.

Moheidat and Baniabdelrahman (2011) in their study stated that the use of Self-Assessment Techniques through the two main tools has revealed many advantages. They noticed that it was very useful especially with shy students. The teachers were able to identify their problems and to get their feedback by using the One -Minute Paper. More importantly, the researchers added that it made students more involved and motivated in the learning process. Finally, they noticed that one more significant advantage to Self-Assessment was that students started to overcome their reading problems and their reading comprehension improved.

SAT and Reading Skills

Baniabdelrahman's study (2010) proved that Self- Assessment was a good technique in EFL instruction. It investigated the effect of Jordanian students' Self-Assessment on their performance in reading in English. The sample consisted of 67 male and 69 female students and was divided into an experimental group and a control group. For the data collection, the researcher used student Self-Assessment tools such as a one-minute paper and rating-scale sheets. The findings of the study revealed that, due to the implementation of the Self-Assessment Technique, there was a significant difference between the reading achievement of the students of the experimental group over the students of the control group. It indicated the positive effect of the Self-Assessment method in improving students' reading comprehension.

Vangah's study (2013) was done with the intention of finding out the effect of Self-Assessment on reading skill and vocabulary Knowledge. The participants were 192 intermediate EFL learners. For collecting the data, the researcher used Oxford placement test, Self-Assessment questionnaire, TOEFL reading comprehension and vocabulary tests. The results showed that Self-Assessment did not have any effect on increase and improvement of vocabulary but it had a great role in reading skill.

El-Koumy (2009) conducted an experiment to investigate the effect of classroom performance assessment on the EFL students' basic and inferential reading skills. The participants consisted of 64 first-year secondary school students in Menouf Secondary School for Boys at Menoufya Directorate of Education (Egypt) during the academic year 2006/2007. The experimental group students were exposed to classroom performance assessment; whereas the control group students were exposed to classroom traditional assessment. The researcher defined the classroom performance assessment as a form of assessment which required students to construct responses rather than select among pre-existing options. Therefore, students in the experimental group used the KWL chart and the Self-Assessment checklist for assessing their own reading strategies and comprehension in each reading session. The KWL chart included Self-Assessment phrases of the lesson such as " What I **Know** about the Topic of the Text, What I **Want** to Know, and What I **Learned** from Reading the Text".

The results showed that there was a statistically significant difference between the two groups in the basic reading skills in favor of the control group and in the inferential reading skills in favor of the experimental group. These findings suggested that classroom performance assessment was less effective in improving secondary school EFL students' basic reading skills, but was more effective in developing their inferential reading skills than traditional assessment.

Previous Studies of SAT in the Saudi Arabian Context

Teaching and learning language skills seem to be benefited by and with Self-Assessment as most studies indicated its positive effect in learners' performance. In KSA, the situation is different as the technique is considered to be less explored especially with college learners. However, the technique has been investigated in schools with intermediate students.

Alabdelwahab (2002) conducted a study with the purpose of examining perceptions of EFL students, EFL teachers and school administrators towards the use of Self-Assessment portfolios. Data were collected from all second-year 81 male students attending the three English as a Foreign Language (EFL) classes at Manarat Al-Sharqiah Intermediate School in the Eastern Province of Saudi Arabia. Those students were asked to use Self-Assessment portfolios as part of their EFL classes during the 16-week Spring semester, 2001. The researcher explained that the Self-Assessment portfolios were designed to enable students to become aware of both successful and unsuccessful strategies they had applied in learning English. Data analysis revealed that 76% of students responded that they enjoyed using Self-Assessment portfolios because portfolios were less stressful than tests and allowed for communication with the teacher.

Also, most students thought that this Self-Assessment approach would be helpful in identifying strengths and weaknesses in their learning. In addition, the study included interviews of two EFL teachers and three school administrators and the results revealed that these individuals endorsed the Self-Assessment portfolio as a type of assessment worthy of future consideration.

Saudi EFL Learners' Difficulties in Learning Reading

According to Alsamadani (2009) there is a huge gap between the world's rapid movement towards teaching reading skills and strategies and the reality of reading instruction in Saudi Arabia. Moreover, Alsamadani (2009) observed a huge gap between Saudi college-level students' proficiency level and that of the higher proficiency expected from EFL university students. This fact is reported in the Educational Testing Services (ETS) of the last years which have shown that Saudi students who have taken the Test of English as a Foreign Language (TOEFL) scored the lowest compared with their peers from Middle Eastern countries (Educational Testing Services, 2013). Thus, in order to improve the teaching of EFL reading in Saudi Arabia, research is needed to understand the learners' reading problems and to evaluate their progress as they progress through EFL programs in the educational system (Alsamadani, 2009). Studies have pointed out that Saudi students in particular are still facing difficulties in learning reading. According to Nezami (2012), the reading skills were not easy in learning for Arabic beginners and university students. In his project, the aim was to find out the comprehension problems in reading faced by EFL Arab learners in general and students of Preparatory Year (boys) and Community College (boys) of Najran University in KSA in particular. Collecting data was on the basis of a survey conducted through the teachers' questionnaire, students' observation, and case studies. The results showed that there were different problems regarding various aspects of reading skills. Finally, the researcher added that "the mother of all problems is that students exhibit very little interest in additional reading materials as they seem unmotivated and uninspired" (p.314). Thus, the researcher argued that EFL learners must be involved in English-language activities. However, it can be said that using SAT could encourage students to be involved in the learning process as shown in the above studies.

In conclusion, with regard to reading skills, Vangah (2013) stated that Self-Assessment Technique and its impacts on reading specially reading comprehension is one of the concerns of teachers in modern approaches of language learning. Therefore, we should consider that Self-Assessment has gained popularity in recent years and its potential value as an instructional tool as well as a measurement tool is recommended.

In the Saudi context, Self-Assessments studies are found in different areas other than language such as medicine, engineering, and science. However, there are almost no studies of the effectiveness of Self- Assessment technique in language skills especially the reading skills. Therefore, it is the aim of this study to find out its benefits for Saudi students.

Methodology

Research Subjects

For the present study, the chosen sample was first-year EFL female students at the Community Service and Applied Studies College in King Saud University. Due to students dropping the class, the sample was reduced to 40 students. The participants were all native speakers of Arabic and had at least 6 years of EFL instruction prior to their admission to the Community Service and Applied Studies College. They were between 18-20 years old and were majoring in translation. The sample could be considered homogeneous because of the students' similar characteristic such as age, gender and educational background. The particular course chosen for this study was the English reading course which was aimed at developing students reading skills. The participants were further divided into an experimental group and a control group of about twenty students each. The control group used only the traditional ways of assessment; on the other hand, the experimental group was exposed to the Self-Assessment Technique.

Research Tools

The research tools were two tools of the Self-Assessment Technique which were adopted from Moheidat and Baniabdelrahman's study (2011): the One- Minute Paper and the Rating- Scale Sheet. They were applied for the experimental group to investigate the technique's effect on the students' performance. The participants in the control group were assessed traditionally by their teacher only.

The One-Minute Paper

It was used at the end of each class period for a few minutes to check the students' understanding of what have been taught in a particular class period. Students had to give short answers to three questions given in the one-minute paper (Appendix A). The questions were:

1. What are the most important things you have learnt in today's class period?
2. Which area(s) of the lesson did you fail to grasp?
3. Which point(s) of the lesson is/are still not clear enough in your mind?

Those questions were significant for both the students and the teacher or the researcher. The first advantage of using that paper was that it helped students to reflect on the reading lessons. They were able to see how they progressed on a daily basis and identified their own problems and the difficulties they still faced. The second advantage of using it was that it gave the teacher immediate feedback about the students' learning and their minor or major problems in each lesson.

The Rating-Scale Sheet

It was used at the end of each chapter to check the students' understanding of the topic as a whole. The textbook had eight chapters and each one had two reading texts and practice or exercises on how to use the three basic strategies of reading. The main idea of teaching the reading strategies was to help students comprehend the reading text, answer the different questions about it and get the best outcome of each text. Therefore, in the rating-scale sheet (Appendix B), students had to check some points related to the basic three strategies of reading which were *Before-I-read strategies, While-I-was-reading strategies* and *After-I-read strategies*

As for the role of the teacher in the Self-Assessment Technique, the data or student feedback from both tools were analyzed and sorted into major topics. Thus, after that analysis, the teacher explained more, added more exercises and tried another teaching strategy for the next class to teach those particular points which students failed to understand.

Research instrument

Pre-Test and Post-Test

For the pre-test, the researcher used a reading test which was used by the English department at the College of Community Service and Applied Studies. The reading test focused on testing the comprehension of a reading passage and the use of the three types of strategies: pre-reading strategies, while reading strategies, and post-reading strategies that would lead to comprehension. The questions included inferring the main ideas, outlining, guessing the meanings of particular words from context, and using different reading strategies. At the end of the study, the same reading test was administered as a post-test to both groups to investigate the students' achievement in reading. The post-test had the same type and number of questions as the pre-test. The pre-test and post-test results were analyzed using the Statistical Package of Social Sciences (SPSS).

Test Validity and Reliability

The reading test was used by the English department at the College of Community Service and Applied Studies in previous semesters.

Thus, the validity of the test format and content has been examined by the head of language department, the translation department coordinator and three experienced EFL teachers working in the College of Community Service and Applied Studies. As for establishing the reliability of the reading test, a pilot study was conducted. The Cronbach's Alpha value for the reliability revealed that the test was reliable at 0.89. Given this result, the instrument was proven to be acceptable for the actual study.

Procedure of the Study

Basically, the experimentation involving the use of Self-Assessment Technique covered a period of two months. During the two months, the technique of student Self-Assessment was used in teaching the subjects of the experimental group to monitor their progress in reading.

The assigned textbook for both groups was *Interaction: Reading Middle East Gold Edition* by Hartmann and Kirn (2007). The textbook had eight chapters and each one had two reading texts and practice and exercises on how to use the three basic strategies of reading. Both groups were taught twice weekly, which covered two hours in each class. The researcher handled the teaching task of the experimental group and the control group was taught by another instructor who relied on the traditional method of assessment only. The whole study lasted 11 weeks with eight weeks allocated for the treatment.

Before conducting the experiment, the selected Self-assessment tools were discussed with the researcher's supervisor to check its suitability to the level of the students and its relevance to the taught skill. In the first week of the semester, the pre-test of reading skills was given to both groups to ensure that the experimental and the control group were equivalent. In the second week, to ensure the appropriate use of the One-Minute Paper, the researcher prepared printed copies of the Self-Assessment tools and trained the subjects of the experimental group before the treatment for two sessions on how to self-assess so that they would be able to use the tools correctly and effectively.

In the third week, the treatment began and it consisted of three basic phases. Firstly, at the end of each class, the One- Minute Paper was distributed while the Rating- Scale Sheet was distributed at the end of each chapter. The researcher gave students time to respond. The students were very enthusiastic about filling the SA sheets. Thus, it took them a lot of time writing exactly what they had learned, opening the book, reviewing the exercises, and trying to be specific about it. Secondly, the researcher went through the students' Self-Assessment tools in order to analyze their feedback. To made it easier, after each class the SA tools of every student are gathered in booklets.

Thirdly, after that analysis, the researcher dealt with students' weaknesses, reinforced their strengths in each class. For example, when the researcher found that the class had failed to understand a certain point; she tried another teaching strategy for the next class to teach that particular point again or added more exercises and explanation. For example, when students mentioned that the major problems for them are getting the topic and the main ideas, the researcher solved it through some ways. She used a lot of exercises using extra materials and gave students those certain points' as homework. Therefore, Not only students had to work on their problems in classroom but they also had to work with it at home as the homework mainly focused on their own difficulties. Then, their answers were corrected and their mistakes were discussed. In addition, in every class the researcher discussed a problem of one student to the whole students. These procedures were repeated throughout the semester for the experimental group specifically from week 3 to 10. In week 11, after the treatment, the post- test of reading skills was given to both groups. The tests scores were analyzed to check if there were any significant differences between the two groups concerning performance in reading skills.

Data Analysis

The results of the pre and post-test were analyzed using the Statistical Package of Social Sciences (SPSS). In statistics, Dornyei (2007) stated that if we want to compare two groups, we compute a't-test'. There were two main types of t-tests used by the researcher to analyze the collected data. Firstly, a t-test for independent means was used. This type of t-test was chosen because the researcher wanted to compare the results of groups that are independent of each other namely the control group and the experimental group (Dornyei, 2007). Thus, it was used to find out whether there were any significant differences between the reading achievements of the experimental group compared to the control group in the pre-test as well as in the post-test. Secondly, a paired –samples t-test was used. This type of t-test was chosen because the researcher aimed at comparing two sets of scores obtained from the same group (Dornyei, 2007). Thus, it was used to find out if there were any improvements in the reading skills of the two groups. In addition, an effect size (ES) analysis was used to determine whether there were improvements in pre-test to post-test scores for the experimental and control groups.

Moreover, it was also used to determine whether there was a difference between the experimental and control groups in the post-test scores of the reading tests.

Results of the Reading Skills

Pre-Test Results of the Reading Skills

To test if there was a significant difference in the level of performance of the experimental group and control group in the pre-test, the t-test for independent means was used. A confidence level of 0.05 was used to determine the statistical significance. For the control group, the reading skills mean and standard deviation were 23.85 and 3.51, respectively. On the other hand, the reading skills mean and standard deviation of the experimental group were 23.85 and 4.29, respectively. Thus, both groups' scores were close, as shown in Table 1. The similarity in the mean scores indicated that the two groups had about the same level of proficiency and language knowledge before undergoing the SAT treatment.

Table 1. Mean Scores and Standard Deviation for the Reading Pre-test

Variable	Group	N	Mean	Std. Deviation
Reading skills pre-test	Control	20	23.85	3.51
	Experimental	20	23.85	4.29

To ensure comparability of the control group and experimental group, a t-test was conducted. The t-test conducted on the pre-test scores of the skills showed t (38) =.000, p=.1.000, α=.05, and this indicated that there was no significant difference between the two groups. In the parametric test (Appendix C), the p (Sig.) values for the reading pre-test across the two groups were greater than 0.05, which was below acceptable confidence levels. Thus, a significant difference between the two groups' results could not be inferred. This ensured that the groups were homogeneous in relation to the level of reading proficiency.

Post-Test Results of the Reading Skills

To answer question one, independent samples t-test was used. After the implementation of the study and on the examination of the data, it was evident that there was an increase in the mean scores in reading skills from 23.85 to 25.90 (difference 2.05) for the control group. Similarly, for the experimental group, the mean scores increased from 23.85 to 27.20 (difference 3.35) (see Table 2). The increase in the mean scores indicated that the two groups had improved at the end of the term and after taking the reading course.

Table 2. Mean Scores and Standard Deviation for Reading Post-Test Scores

Variable	Group	N	Mean	Std. Deviation
Reading Skills	Control	20	25.90	2.71
	Experimental	20	27.20	2.91

Another t-test was conducted to determine if the post-test scores of the reading skills for the groups were significantly different. In the t-test of the reading skills, the t value was -1.461 and the significance level was p=.152 (Appendix C). The results (t (38) =-1.461, p=.152, α=.05) indicated that there was no significant difference in the performance of the two groups on the reading post-test. Because the significance value of .152 was greater than alpha=0.05, we could conclude that there was no significant difference between the groups in the reading post-test. In the parametric test (Appendix C), the p (Sig.) values for the reading post-test across the two groups were greater than 0.05, which was, falling below acceptable confidence levels. Thus, no significant difference between the two groups' results could be inferred.

In addition to the test of significance, an effect size (ES) analysis was used. The effect size was another analysis of the statistical results of the study to determine the effect of the Self-Assessment Technique and to measure its strength. One way to interpret the effect size or to measure the effect size was d, known as Cohen's d (1988). In the present study, firstly, an effect size analysis using Cohen's d (1988) was used to determine whether there was a difference between the experimental and control groups on the post-test scores of the reading skills. For the reading skills post-test, we had an effect size of d=0.46. This indicated that the post-test mean of the experimental group was at the 66th percentile of the control group. The experimental group showed slightly higher reading skills based on their post-test scores than the control group. According to Cohen (1988), the difference between the two groups was considered medium. Therefore, the effect of the Self-Assessment Technique was not large.

Post-Test Improvement in the Reading Skills

To answer question two, paired samples t-test was used. The paired samples t-test on the pre-test and post-test scores of the control and experimental groups was conducted to find out if there were any improvements in the reading skills of the two groups.

For the control group, the paired sample t-test revealed a statistically significant difference between the mean scores of reading pre-test (mean=23.85, sd= 3.51) and post-test (means= 25.90, sd=2.71) and that the results showed t (19) =-2.483, p=.023, α=.05 (Appendix D). Because the significance value of .023 was smaller than alpha=0.05, we could conclude that there was a significant difference between the pre-test and post-test of the control group.

For the experimental group, the paired sample t-test also revealed a statistically significant difference between the mean scores of reading pre-test (mean=23.85, sd=4.29) and post-test (mean=27.20, sd=2.91) and that the results showed t (19) = -4.264, p=.000, α=.05 (Appendix E). Because the significance value of .000 was smaller than alpha=0.05, we could conclude that there was a significant difference between the pre-test and post-test of the experimental group.

Thus, in terms of statistical meaning, although both groups showed overall improvements in their test scores, there was no significant difference between the two groups in the post-test.

In addition, an effect size analysis was used to determine whether there were improvements in pre-test to post-test scores for the experimental and control groups. For the reading skills test of the control group, we had an effect size of d= 0.65. Cohen (1988) would describe this effect size as 'medium'. Thus, the students in this group showed medium improvement in reading skills scores from pre-test to post-test. On the other hand, for the reading skills test of the experimental group, we had an effect size of d= 0.91. Cohen (1988) would describe this effect size as 'large'. Thus, the experimental students showed large improvement in reading skills scores from pre-test to post-test.

The Self-Assessment Sheets

One-Minute Paper

The one-minute paper was used to check students' understanding of what had been taught in every class period. For example, some students stated that they had learned using the parts of speech to understand vocabulary. Others declared that they had learned how to get meaning from context by understanding certain abbreviations. On the other hand, most students revealed that they had learned how to use some reading strategies especially the *Before-I-read strategies*. By the end of the term, students began to mention learning how to identify the topic and the main idea of a given paragraph and answering the comprehension questions of a given text. Moreover, they assessed learning how to use some reading strategies specifically the *While-I-was-reading strategies* and the *After-I-read-strategies*.

On the other hand, for the difficulties, some students stated that they faced problems with guessing meaning of new vocabulary from context. Others said that they encountered difficulty in finding the topic and the main idea of a given paragraph. A third group mentioned that they had problems with comprehension exercises while others encountered difficulty with language exercises such as identifying suffixes and prefixes of the texts. Basically, the SAT tool showed that students responded well to the question sheet.

Rating-Scale Sheet

The second tool was used at the end of each chapter to check the students' understanding of the topic as a whole. Students had to check some points related to the basic three strategies of reading. The researcher noticed that students' responses varied during the study. For example, during the coverage of the first few chapters of the textbook, most students ticked YES on using the *Before-I-read strategies*. On the other hand, for the *While-I-was-reading strategies* and *After-I-read strategies,* most students ticked NO which suggested that they either had problems with the strategies or they did not know how to use them. For the middle and last chapters of the book, students began to show some progress. For example, they ticked YES on using some of the *While-I-was-reading strategies* and *After-I-read strategies.* Thus, the main purpose of the instrument was to make the students aware of these strategies.

Discussion

Research Hypothesis 1

The present study hypothesized that the Self-Assessment Technique does not affect the performance of the experimental group in reading skills. After conducting the experiment, the results indicated that there was not a statistically significant difference at the 0.05 level of significance between both the experimental and the control groups in the reading post-test. Therefore, the first null hypothesis was accepted. This result might be due to several reasons. First, the students' lack of Self-Assessment practice and training might have kept them from getting the most benefit from the SAT. In the study, the students of the experimental group were trained before the treatment for only two sessions. Second, students' unawareness of the advantages of Self-Assessment might have led them to not participate fully in assessing their own reading skills. Third, the short time of the study might have decreased the benefits of SAT for the experimental group. The whole study lasted 11 weeks. However, due to sessions which had been given to the training, pre-test and post-test, the experimental group had only eight weeks allocated for the treatment.

Fourth, the small number of the study sample with only 20 female students in the experimental group was not enough for a final judgment regarding the value of SAT. Furthermore, the difficulty of acquiring the reading skills by Saudi students might have kept the SAT from arriving at its goals.

Results of this study were in agreement with El-Koumy (2001) whose study was done to investigate the effects of Self-Assessment on EFL students' knowledge achievement and academic thinking. It revealed that the difference between the groups was not significant at the 0.05 level. Based on that, we could state that SAT does not affect the performance of EFL learners. On the other hand, the results of the present study did not coincide with some other studies in which the effect was found for the Self-Assessment Technique in reading skills such as Baniabdelrahman's (2010) Vangah 's (2013) and El-Koumy 's (2009) studies. Furthermore, those studies were conducted on a larger sample than the present study.

Moreover, the importance of time for SAT to be used well was stressed by Oscarson (2009) while the importance of practice and training of students on how to assess was emphasized by Munoz and Alvarez (2007) and Ross (2006). Finally, the difficulty of acquiring the reading skills by Saudi students might have kept SAT from achieving its goals. This was in accordance with the overall findings from some Saudi researchers about the low readings skills of Saudi students such as Al-Nujaidi (2003). In addition, Alsamadani (2009) noticed the huge gap between Saudi college-level students' proficiency level and that of the higher proficiency expected from EFL university students. As I have mentioned earlier, according to Nezami (2012), learning reading skills were not easy for Arabic beginners of English and university students.

Research Hypothesis 2

The present study hypothesized that there is no significant improvement in the reading skills of the two groups. After conducting the experiment, the results indicated that there was a statistically significant difference at the 0.05 level of significance between the pre-test and post-test scores in the reading skills of the two groups in favor of the post-test scores. Thus, the findings rejected the null hypothesis of the study. This finding could be explained as follows: firstly, at the time of the pre-test, students in both groups were not taught the reading course and thus they were with insufficient practice and were not familiar with the use of reading strategies. Secondly, at the time of the post-test, students in both groups were taught the reading course and therefore they were with sufficient practice. Also, they were familiar with the use of reading strategies.

Results of this study were in agreement with some studies of SAT in language learning on which both groups significantly improved. For example, Meihami and Varmaghani's (2013) investigation of SAT in writing found that participants in both groups significantly improved in writing. Also, Baniabdelrahman's (2010) study of SAT in reading has proved that both groups improved. Moreover, El-Koumy's (2009) study of SAT in reading showed that both groups improved differently. In addition, the present study's result was in accordance with the overall findings of some Saudi researchers in the role of reading strategies in improving students' reading performance. Similarly, Alsamadani's (2011) study showed that using the different reading strategies was important to improve Saudi learners' comprehension. On the other hand, Al-Nujaidi (2003) found that female students in particular used those reading strategies more than male students. Therefore, using the taught reading strategies helped in improving Saudi reading skills.

Conclusion

Focusing on female EFL learners enrolled in the English language translation program at female Community Service and Applied Studies College in King Saud University in Riyadh, the present study attempted to examine the effectiveness of Self-Assessment as a teaching technique on the reading skills of Saudi college students. The study had two groups (control vs. experimental) and the participants included 20 first-year female Arabic students selected to be the subjects of the experiment. The study employed for the pre-test and post-test a reading test which was used by the English department at the College of Community Service and Applied Studies. The Reading test focused on testing reading skills such as comprehension and strategies which were the core of the reading course taught for both groups in this study. For the SAT, the study employed two types of SAT tool which were the One- Minute Paper (Appendix A) and the Rating- Scale Sheet (Appendix B).

To answer the research questions and test the hypotheses, the analyses of the pre-test and post-test data were conducted using SPSS for windows. In addition, descriptive statistics were used including means and standards deviations. The findings of the reading post-test scores of both groups revealed that the mean scores of the experimental group after being involved in SAT were higher than those of the control group; however, the t-test for independent means showed that there was no statistical significant difference between both groups. In addition, in the ES analysis the experimental group showed slightly higher reading skills post-test scores than the control group.

Moreover, the results showed that the mean scores of both groups on the post-test were higher than their mean scores on the pre-test. Furthermore, it was revealed that there was a statistically significant difference in favor of the post-test scores for both groups. In addition, in the ES analysis, the experimental students showed large improvement in reading skills test scores from pre to post while the control students showed medium improvement in reading skills test scores from pre to post.

In conclusion, the results in general did not support the beneficial effects of Self-Assessment as a powerful teaching technique in the improvement of EFL learners reading skills, as the students in the experimental group did not outperform those in the control group in terms of their reading skills.

Implications of the Study

Based on the results, some pedagogical implications are concluded:

1. Self-Assessment could be a very useful tool for teachers. It provides teachers with useful clues and notes about students' weaknesses and strengths. In addition, the technique encourages teachers to change their own teaching strategies and constantly looking for the best approach to fit students' needs.

2. For Self-Assessment to achieve its full potential, teachers should give students enough time to get used to the technique, explain its goals and give sufficient training sessions.

3. Workshops and training courses should be offered by experts within the field of SAT to help language teachers and guide them in using this technique.

4. Teachers are recommended to design SAT tools for language skills and components.

Suggestions and Recommendations for Further Research

Based on the results of the present study, the following are recommended:

1. The study investigates the effectiveness of SAT on reading. Further studies are recommended to investigate SAT on other language skills and components.

2. It is recommended to investigate SAT on gender and other learning variables.

3. It is recommended to implement a similar study taking teachers' attitudes into consideration.

4. A replication of this study with bigger sample is recommended to see if the results will be repeated or changed.

5. This study is limited to approximately one semester of learning. Therefore, more time is necessary to investigate the impact of SAT in reading skills.

References

Alabdelwahab, S. Q. (2002). *Portfolio assessment: A qualitative investigation of portfolio self-assessment practices in an intermediate EFL classroom, Saudi Arabia* (Doctoral dissertation) (Order No. 3081894). Retrieved from ProQuest Dissertations & Theses database. (305587614)

Al-Nujaidi, A. H. (2003). *The relationship between vocabulary size, reading strategies, and reading comprehension of EFL learners in Saudi Arabia* (Doctoral dissertation) (Order No. 3094023). Retrieved from ProQuest Dissertations & Theses database. (305242924)

Alsamadani, H. A. (2009). *The relationship between Saudi EFL college-level students' use of reading strategies and their EFL reading comprehension* (Doctoral dissertation) (Order No. 3353336). Retrieved from ProQuest Dissertations & Theses database. (304963586)

Alsamadani, H.A. (2011). The effects of the 3-2-1 reading strategy on EFL reading comprehension. *English Language Teaching*, *4*(3), 184-191. doi:10.5539/ELT.V4N3P184

Baniabdelrahman, A. A. (2010). The effect of the use of self-assessment on EFL students' performance in reading comprehension in English. *The Electronic Journal for English as a Second Language, 14*(2), 1-22. Retrieved from
http://www.tesl-ej.org/wordpress/issues/volume14/ej54/ej54a2/

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Dornyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative, and mixed methodologies*. Oxford: Oxford University Press.

Educational Testing Services (2013). *Test and score data summary for TOEFL internet-based and paper-based tests*. Retrieved March 23, 2014, from http://www.ets.org/toefl

El-Koumy, A. S. (2001*). Effects of student self-assessment on knowledge achievement and academic

thinking*. Retrieved from http://files.eric.ed.gov/fulltext/ED452731.pdf

El-Koumy, A. S. (2009). *The effect of classroom performance assessment on EFL students' basic and*

inferential reading skills. Retrieved from http://files.eric.ed.gov/fulltext/ED514530.pdf

Harris, M., & McCann, P. (1994). *Assessment: Handbooks for the English classroom.* Oxford:

Heinemann.

Harris, M. (1997). Self-assessment of language learning in formal settings. *ELT Journal, 51*(1), 12-

20. doi:10.1093/ELT/51.1.12

Klingner, J., Vaughn, S., & Boardman, A. (2007). *Teaching Reading comprehension to students with*

learning difficulties. New York and London: Guilford press.

Marhaeni, A. A. (2005). Self-assessment in EFL instruction: Why does it matter? Retrieved from

http://pasca.undiksha.ac.id/e-learning/staff/images/img_info/4/17-282.pdf

Meihami, H., & Varmaghani, Z. (2013). The implementation of self-assessment in EFL writing

classroom: An experimental study. *International Letters of Social and Humanistic Sciences, 9,*

39-48. Retrieved from

http://www.ilshs.pl/wp-content/uploads/2012/11/ILSHS-9-2013-39-481.pdf

Moheidat, A. S., & Baniabdelrahman, A. A. (2011). The impact of Omani twelfth-grade students'

self- assessment on their performance in reading in English. *Asian EFL Journal, 13*(1), 48-84.

Retrieved from http://www.asian-efl- journal.com/PDF/March-2011-asm.pdf

Munoz, A., & Alvarez, M. E. (2007). Students' objectivity and perception of self-assessment in an

EFL classroom. *The Journal of Asia TEFL, 4*(2), 1-25. Retrieved from

http://www.eafit.edu.co/idiomas/centro-

idiomas/Documents/Student%C2%B4s%20Objectivity%20and%20Perception%20of%20

Self%20Assessment%20in%20an%20EFL%20Classroom.pdf

Nezami, S. R. (2012). A critical study of comprehension strategies and general problems in reading

skill faced by Arab EFL learners with special reference to Najran university in Saudi Arabia.

International Journal of Social Sciences and Education, 2(3), 306-316. Retrieved from

http://ijsse.com/sites/default/files/issues/2012/Volume%202%20Issue%203,%202012/Paper-

30/Paper-30.pdf

Oscarson, M. (1989). Self-assessment of language proficiency: Rationale and applications. *Language

Testing, 6*(1), 1-13. doi:10.1177/026553228900600103

Oscarsson, M. (1997). Self-assessment of foreign and second language proficiency. In C. Clapham

and D. Corson (Eds.), *Encyclopedia of Language and Education, Volume 7: Language

Testing and Assessment* (pp. 175-187). Netherlands: Kluwer Academic Publishers.

Oscarson, A.D. (2009). *Self-assessment of writing in learning English as a foreign language: A study

at the upper secondary school level* (Unpublished doctoral dissertation). University of

Gothenburg, Sweden.

Ross, J. A. (2006). The Reliability, validity, and utility of self-assessment. *Practical Assessment,

Research & Evaluation, 11*(10), 1-13. Retrieved from

http://pareonline.net/getvn.asp?v=11&n=10

Vangah, F.P. (2013). Effect of self-assessment on Iranian EFL learners' reading skill and vocabulary

knowledge. *International Research Journal of Applied and Basic Sciences, 4*(3), 676-680.

Retrieved from http://www.irjabs.com/files_site/paperlist/r_731_130328110828.pdf

APPENDICES

APPENDIX A

One-Minute Paper (English Version)

Dear Student,

Self-assessment is a vital component in learning. You are kindly requested to comment frankly on your own learning. You are encouraged to share in the responsibility for your own learning. Teaching cannot be effective unless the teacher comes to know your strengths and weaknesses. Your responses will help me find out how the course is going and give me an idea whether any changes are needed. This one-minute paper is mainly prepared to help you, and it won't take much of your time. The more you are open and accurate in answering these questions, the more progress in teaching and learning there will be.

1. What are the most important things you have learnt in today's class period?

………………………………………………………………………………
…………………………………………………………………………

2. Which area(s) of the lesson did you fail to grasp?

………………………………………………………………………………
…………………………………………………………………………

3. Which point(s) of the lesson is/are still not clear enough in your mind?

………………………………………………………………………………
…………………………………………………………………………

I appreciate your cooperation.

Adopted from Moheidat & Baniabdelrahman (2011).

APPENDIX B

Rating-Scale Sheet (Self-assessment for Reading)

Chapter (   )

Name:	**Date**:......../ ..../..........

Before I read "…………………," I:

Yes   No

_____thought about the title and what it suggested the text was about.

_____previewed the whole text or parts of it.

_____thought about the subject or situation.

_____set a purpose for my reading.

While I was reading "…………………………………," I:

Yes   No

_____developed a dialogue with the writer (e.g., what is the writer
communicating? what is the main idea? what do I already know about this?).

_____visualized what places, people, and events might look like.

_____connected my personal experience to what I was reading.

_____made inferences from textual clues given by the writer.

_____tried to distinguish between fact and opinion.

_____predicted and then checked what the writer might say next.

_____went over the parts I found confusing.

_____checked words that I did not know the meaning of from context.

After I read "…………………………………," I:

Yes   No

_____determined an initial impression of what I had read.

_____discussed what I had read and my impressions with someone.

_____reflected on what I had read.

_____reviewed and summarized what I had read and learned.

_____made notes in my notebook.

_____developed a more thoughtful interpretation of what I had read
(considered why the writer wrote the text, what was being presented, and
how it was constructed).

_____evaluated what I had read and supported my judgments with references
to the text.

Adopted from Moheidat & Baniabdelrahman (2011).

APPENDIX C

Independent Samples t-Test for Pre- and Post-test Scores of Two Groups

Group Statistics

Group		N	Mean	Std. Deviation	Std. Error Mean
Pre-Reading Test	Control	20	23.8500	3.51351	.78564
	Experimental	20	23.8500	4.29535	.96047
Post-Reading Test	Control	20	25.9000	2.71254	.60654
	Experimental	20	27.2000	2.91277	.65131

Independent Sample Test

		Levene's Test for Equality of Variances		t-Test for Equality of Means							
		F	Sig.	t	df.	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Pre Reading Test	Equal variances assumed	.838	.366	.000	38	1.000	.0000	1.24086	-2.51199	-2.51199
	Equal variances not assumed			.000	36.563	1.000	.0000	1.24086	-2.51524	-2.51524
Post Reading Test	Equal variances assumed	.000	1.000	-1.461	38	.152	-1.3000	.89000	-3.10172	.50172
	Equal variances not assumed			-1.461	37.809	.152	-1.3000	.89000	-3.10202	.50202

APPENDIX D

Paired Samples t-Test for Pre- and Post-test Scores of the Control Group

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	PRE-Reading Test	23.8500	20	3.51351	.78564
	POST-Reading Test	25.9000	20	2.71254	.60654

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	PRE-POST Reading Tests	20	.319	.171

Paired Samples Test

		Paired Differences					t	df.	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	PRE-POST Reading Tests	-2.0500	3.69174	.82550	-3.7778	-.3222	-2.483	19	.023

APPENDIX E

Paired Samples t-Test for Pre- and Post-test Scores of the Experimental Group

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	PRE- Reading Test	23.8500	20	4.29535	.96047
	POST- Reading Test	27.2000	20	2.91277	.65131

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	PRE-POST Reading Tests	20	.583	.007

Paired Samples Test

		Paired Differences					t	df.	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	PRE-POST Reading Tests	-3.3500	3.51351	.78564	-4.9944	-1.7056	-4.264	19	.000