Task Based English Language Teaching in Saudi Intermediate Schools

Sultan A. Al Muhaimeed

dr.s.almuhaimeed@gmail.com

Qassim University

Saudi Arabia

This study investigates the effectiveness and appropriateness of a constructivist instructional practice for EFL. It strives to determine whether adopting Task-Based Language Teaching (TBLT) is a more effective means to increase students' reading comprehension when compared to that of the traditional teaching method. It also attempts to gain understandings that accompany TBLT implementation through constant comparison and contrast them with those that accompany the traditional teaching method. The mixed-method study covers quasi-experimental approach that uses one pretest and several posttests to collect quantitative data, as well as classroom observation and researcher log to collect qualitative data. The independent variable is the use of TBLT and the dependent variable is the students' reading comprehension achievement scores. A Two-Factor Split Plot analysis with pretest as the covariate is used for analyzing the quantitative data. The analysis of qualitative data includes synthesis, rich, and detailed descriptions for classroom observation and grounded theory for researcher log data.

Keywords: Task Based Language Teaching, Constructivist Instructional Practices

Introduction

The employment of tasks was initiated in the field of Second Language Acquisition (SLA) towards the end of 1960s and at the beginning of 1970s (Burt & Dulay, 1973; Hakuta, 1976; Krashen, 1994; Long, 1996). Tasks, at the beginning, were employed to describe particular aspects of language acquisition such as that of grammar and, later, were based on theories such as those related to language production. Across time, the term "task" was sometimes used as a synonym of problem-solving and role-plays techniques and vice versa (Brown, 2007). The use of tasks in English language teaching and learning was linked to the development of SLA research (Ellis, 2003).

Literature shows that tasks in the learning processes are used as means of assessment or as a method of teaching and learning. The later one is the focus of the present study and holds the terms Task Based Language Teaching TBLT. TBLT has developed over time until it has reached the formation and complexity this study has adopted. TBLT is a method of language teaching in which meaning is primary, there are real world problems to solve, and priority is placed on the completion of the tasks, which are assessed in terms of the outcome (Brown, 2007; Willis & Willis, 2007). A task-based learning situation is organized in accordance with the three stages of a task pre-task, during task, and post-task (Ellis, 2003). This method is seen by many researchers to be emerging from Communicative Language Teaching (CLT) (Brown, 2007; Ellis, 2003). Others see it as a new approach to English language teaching and learning (Kumaravadivelu, 2006).

Literature Review

There is a new social attitude that argues that there is no one comprehensive theory for learning but, instead, a combination of learning theories or inter-disciplinary learning theories (Jarvis, 2006; Jarvis & Parker, 2005). TBLT is grounded in constructivist theories, which adds strength and value to this method of teaching. The following intends to demonstrate how constructivist theories embody TBLT.

TBLT is theoretically framed by Piagetian (cognitive) and Vygotskian (sociocultural) perspectives as they both emphasize the role of social interaction in cognitive development (Piaget, 1970; Vygotsky, 1978) and which is a fundamental component of TBLT (Lee, 2000). For clarification, the cognitive and sociocultural perspectives of learning, the psychological (cognitive) theories trace the arrows from the person to the external objectified culture while sociological (sociocultural) learning theories start with objectified culture and point inwards to the individual and, hence, learning should be seen from both perspectives (Jarvis, 2009).

Principles of TBLT suggest that it is theoretically framed and influenced by Vygotsky's (1978) sociocultural theory. One Principle is that learning through interaction among learners is a fundamental principle of TBLT (Lee, 2000). At the same time, this principle is partially resembled by or linked to mediation in the sociocultural theory and which, in essence, suggests that learning is socially constructed (Vygotsky, 1978). Another principle is that tasks in TBLT include mediation by others in social interaction, by self through private speech, and through artifacts (Lantolf, 2000).

From a cognitive perspective, knowledge is the product of learning and is neither totally external nor totally internal, but a result of interaction between heredity (internal) and environment (external) (Piaget, 1969). This theoretical perspective embodies the TBLT method involves two factors to be present when producing the language, which are (a), the simultaneity of the information processed by the learner and (b) involvement in context (Cummins, 1983). In other words, the nature of tasks requires students to have a reciprocal interaction of language with their colleagues through production (within the self) and reception (from the environment).

Tasks Foundational Thinkers across Time

Defining and theorizing tasks have developed and accumulated across time beginning in the mid of 1980s (Breen, 1989; Bygate, Skehan, & Swain, 2001; Crookes, 1986; Ellis, 2003; Lee, 2000; Long, 1985; Nunan, 1989; Prabhu, 1987; Richards, Platt, & Weber, 1985). These attempts have sometime provided literature with a broader definition of tasks (Breen, 1989; Crookes, 1986; Richards, et al) and sometime with a narrower definition of tasks (Bygate, Skehan, & Swain, 2001; Ellis, 2003; Lee, 2000; Long, 1985; Nunan, 1989; Prabhu, 1987). The broader attempts have suggested that a task is an activity that helps accomplish language learning or simply a piece of work that provides learners with opportunity and knowledge to communicate in the target language. The narrower attempts have suggested that a task is a piece of work related to the real world, facilitated by the teacher, urged learners to comprehend, manipulate, produce, and interact in the target language, and call for primarily meaning-focused language use.

Examining the definitions of tasks by separately each foundational thinker shows that the definition and design of the tasks for this study are based on two perspectives. The first suggests that a task is an activity characterized by interaction, structuring and sequencing, focus on meaning, comprehension, manipulation, and production of the target language (Lee, 2000). The second suggests that a task requires learners to use the language in a meaningful way (Bygate, Skehan, and Swain, 2001).

Research on Task-Based Learning and Teaching

Research literature has addressed important aspects or issues of task-based learning and teaching (De Bot, 2001; Kim, 2008; Rivers, 2010; Skehan, 1998; Stevens, 1983; Swain & Lapkin, 2000; & Swan 2005). These aspects included the complexity of tasks, efficiency of task-based learning and teaching, enhancing language production when adopting task-based learning, and criticism of task-based learning and teaching. The following presents the reviewed research literature on each of these aspects.

Complexity of tasks. The complexity of tasks has been a central theme for task-based research due to its immediate relevance to learner production (Carless, 2008; Robinson, 2001; Skehan, 1998). The (inner) complexity of tasks influences learner production and, hence, attention needs to be paid for sequencing tasks on the basis of their cognitive complexity is preferable to sequencing them based on difficulty (the learner's building new knowledge on previous obtained or learned knowledge). It is obvious that emphasis on the (inner) complexity of tasks goes along with the principles of the Zone Proximal Development ZPD proposed by Vygotsky (1978). The (outer) complexity of tasks needs to be organized and designed carefully due to its beneficial effects on learner production. Task-based lesson is organized in accordance with the three stages of a task (pre-task, during task, and post-task) (Ellis, 2003).

Efficiency of tasks. Learning efficiency through tasks has been demonstrated in research literature of task-based learning and teaching (Brown, 2007; Swain & Lapkin, 2000). Task-based teaching promotes pedagogical tasks which form nucleus of the classroom activity. Also, Oral and written tasks provided the learners with opportunities to learn language. In addition, children know far more language through activities (tasks) than what they exhibit in response to classroom drills (Stevens, 1983). Moreover, artifacts such as tasks helped in tracking learner development over time and also contributed to shaping the teacher's interactions with learners as they pulled into focus a range of mediating alternatives of varying explicitness (Poehner, 2009).

Research on efficiency of tasks in teaching extends to suggest that tasks participate in creating a real purpose for language use and provides a natural context for language study (Izadpanah, 2010; Swain & Lapkin, 1998). In other words, learning through tasks helped students learn language since the context the tasks present does provide the students with a real learning purpose. Providing learners with a purpose while performing tasks helped students also solve linguistic problems through dialogue. At last, efficiency of tasks showed a significant effect on acquisition in experimental settings (De Bot, 2001; Swain & Lapkin, 2000).

Enhancing TBLT. Another aspect of TBLT research literature is related to enhancing language production when adopting task-based learning in classroom (Albert & Kormos, 2004; Kim, 2008). For instance, creativity is found to affect participants' output in oral narrative tasks moderately. Similarly, learner's higher level of involvement during the task promoted more effective initial vocabulary learning and better retention of the new words. Also, learners need to know that the task outcome (completing the task) is the most important thing as the purpose of the task to use the language rather than display it (Ellis, 2003). In addition, learning through any given task should represent a rehearsal for future social or professional interaction (Rivers, 2010).

Criticism of TBLT. However, fewer researchers had quite different opinions or were skeptical about the efficiency of tasks in language teaching (Mohamed, 2004; Swan 2005). For instance, task-based instruction is not greatly better than the traditional methods and is based on unproved hypotheses (Swan, 2005). Also, learners' preferences relating to deductive and inductive tasks and how learners see the effectiveness of both types showed that learners see both types to be useful and there are no significant differences in learners' preference to either type (Mohamed, 2004). Although those opposing perspectives about the efficiency of tasks in language teaching were explained in TBLT literature as discussed below, the main purpose that made this study bring those two opposing perspectives prior to the application of this study is to have an outlook to which perspective the results of this study might support.

Criticizing perspectives of the efficiency of tasks in language teaching (those who are not in favor of or against task-based learning) are also explained or challenged in TBLT research literature (Murphy, 2003; Plews & Zhao, 2010; Swain and Lapkin, 2000). One explanation in a study about Canadian English as a second language revealed that teachers adapt TBLT in ways that do not go along with or contradict its theoretical principles. In other words, some teachers implement TBLT in way that they do not follow all of TBLT principles and, hence, students might not get the desired outcome. This makes teachers mistakenly refer this problem to the application of TBLT. Another explanation of task-based learning lacking significance is attributed to factors beyond its principles such as that of learners' influence that is found to jeopardize or hinder the task designer's goals. A final explanation for the emergence of the opposing perspectives about the efficiency of tasks in language teaching could be attributed to the need of further research to cover various aspects of task-based learning and teaching (Candlin, 2001; Samuda & Bygate, 2008) and one of the purposes of this study is to participate in adding a line to the TBLT literature.

Purpose of the Study

This study participates in the efforts invested to overcome the dissatisfaction about the students' English language proficiency in the Saudi Arabian context (Maroun & Samman, 2008). It is based on the argument that the lack of sufficient English language proficiency, partially, is associated with the existence of the traditional way of English language teaching in schools. This traditional way of teaching includes instructional practices that are collectively referred to as 'prompting' because they involve the prompting and 'drilling' of students. Due to research scope purposes, this study proposes that the traditional way of English language teaching does not help students better comprehend English when the emphasis is placed on reading comprehension. In contrast to this behaviorist approach, this

study hypothesizes that adopting a constructivist instructional practice (Task-Based Language Teaching TBLT) might lead to improved language reading comprehension.

Research Questions

1. Is using the TBLT method for teaching English as a second language for male third-grade students in intermediate schools in Saudi Arabia more effective in the acquisition of the English language, in terms of students' achievement on reading comprehension, than using the traditional "prompting" method?
2. What insights and issues can be gained about implementing TBLT in this research setting?

Methodology

Design and Method

The study is based on a mixed method design (quantitative and qualitative) where the quantitative part includes a two-factor split-plot analysis with a pretest (covariate) and posttests as a part of quasi-experimental design. The qualitative part is based on observational data and a researcher log.

Table 1.

Two-factor Split-plot Analysis

Factor 1B		Factor 1A		
Traditional teaching method group (Control group)		**TBLT teaching method group (Treatment group)**		
O		**O**		
CT		**XE**		Weeks 1 & 2
OTET	ORPA	OTET	ORPA	Section A
ORPA	OTET	ORPA	OTET	Section B
CT		**XE**		Weeks 3 & 4
ORPA	OTET	ORPA	OTET	Section A
OTET	ORPA	OTET	ORPA	Section B
CT		**XE**		Weeks 5 & 6
OTET	ORPA	OTET	ORPA	Section A
ORPA	OTET	ORPA	OTET	Section B
CT		**XE**		Weeks 7 & 8
ORPA	OTET	ORPA	OTET	Section A
ORPA	ORPA	OTET	ORPA	Section B
CT		**XE**		Weeks 9 & 10
OTET	ORPA	OTET	ORPA	Section A
ORPA	OTET	ORPA	OTET	Section B

O = Pretest
XE = Experiment (treatment) group
CT = Traditional (control) group
ORPA = Observation (researcher prepared assessment posttest)
OTET = Observation (textbook established posttest)
Sections A & B = both of the treatment and control groups have two sections

Researchers who employ quasi-experimental designs rely on various techniques to control (or at least reduce) the threats to the internal validity of the study. In this study, one technique is to randomly assign the classrooms to the treatment and control groups (Wiersma & Jurs, 2009). Variables such as the students' gender, age, and citizenship, the time of the class, classroom settings, teaching aids, the teachers, and the school are already being controlled for due to the design of the study or statistically in the analysis of the study (see Figure 1).

Procedures and Data Collection Tools

To gather data, the study uses: (a) a pretest—to document the level of students' English language reading comprehension they have at the beginning of the study; (b) observation of the treatment group (researcher log) and control group (classroom visits); and (c) posttests—to evaluate particular areas of student study.

Pretest. The primary purpose of the pretest is to function as the main covariate. It is used to provide a baseline for the students' current English language reading comprehension levels so this study can examine the effect of treatment, relative to initial English language proficiency. The pretest also helps increase the power of the study by reducing the error that can be attributed to prior differences among students and its relation to the outcome. It is not used to place students in certain levels or groups. Two reading passage practice tests were used and which were developed by Ohio Department of Education as the pretest in this study. The two reading passages have a total of 22 questions that measure reading comprehension. Test scores are going to be based on a retelling rubric.

Observation. This study places great emphasis on this data collection tool and is aware that field notes gathered are going to represent the eyes, ears, and the perceptual senses of the reader (Patton, 2002).

Table 2.

Classroom Observation Checklist

	Response			
Elements	Yes	No	N/A	Comments
Setting the stage	□	□	□	
Task sequence	□	□	□	
Engaging the students	□	□	□	
Running the task	□	□	□	
Task completion	□	□	□	
Students' attitudes	□	□	□	
Teacher's attitude	□	□	□	
Difficulties	□	□	□	
Advantages	□	□	□	
Disadvantages	□	□	□	

The form and notes provides insights and issues about implementing both of the TBLT (gathered by a knowledgeable colleague) and traditional (gathered by the researcher) methods in the control and treatment groups in this research setting. Observational data are accurate, detailed, and rich in nature (Schram, 2006). Among the techniques used when taking field

notes is the usage of direct quotes, paraphrases, description of the context, and description of any behavioral experience that take place in the classroom (Schneider, 2005).

Researcher log. The treatment teacher (researcher) creates a log where he writes down notes as he recalls them by the end of each day he teaches the treatment group (McNiff, Lomax, & Whitehead, 1996). Those notes include insights and issues about implementing TBLT in this research setting. Out of the numerous visits of writing to researcher log, data can be described by being rich, detailed, and accurate (Schram, 2006).

Recalled data in researcher log can take the form of direct quotes, paraphrased responses, conversations, description of the context, and description of any behavioral experience that take place in the classroom (Schneider, 2005). When logging into researcher log by the end of each day, recalled data would be related to three types of interactions that take place in classrooms: student-to-student (such as group work or pair activities), student-to-teacher interaction (such as instruction by the teacher or questions and participation by the students), and students-to-curricular materials (such as textbooks and workbooks).

Posttests. Posttests are administered at the end of each two weeks to assess students' reading comprehension on the content covered in those two weeks. This process continues for ten weeks. Each posttest consists of two formats; one is the researcher's prepared assessment (RPA) and the other one is the text established test (TET) and which students need to do it all in English. The RPA posttest is mainly retelling where the students read a passage and then are asked to retell the passage using on their own words as they have understood it. The second posttest format, text established test (TET), is an existing instrument provided by the curriculum.

Retelling rubric. Test scores of both the pretest and posttests are based on a retelling rubric originally designed by Applegate, Quinn, and Applegate (2008) and developed in this study as shown in Table 3.

Table 3.

Retelling Rubric

Score	Description
4.0	A virtually perfect retelling that includes all information and a well-supported personal response
3.0	A very strong retelling that includes all information, but does not include a personal response
2.0	A solid retelling that includes most information but that is also characterized by some key omissions and that may include a personal response
1.0	A weak retelling that includes little information but is also characterized by some glaring omissions and factual distortions and that does not include a personal response
0	A retelling that may include nothing more than a vague idea of the topic of the text and that does not include a personal response

The retelling rubric consists of five grading levels ranging from zero to four where four represents the highest grading score and zero represents the lowest one.

Internal & External Validity

To establish internal validity, the relationship observed between two or more variables should be unambiguous and not attributable to something else (Fraenkel, Wallen, & Hyun, 2012). Many of the possible threats to the internal validity of the study are controlled by its design. The study has taken steps to maintain internal validity—for example; the subjects' characteristics are captured and controlled through the pretest, which provides a base line to eliminate the threat of subjects' characteristics to the internal validity.

As for external validity, the processes and the tests are the same country-wide, the findings of this study are potentially relevant for all intermediate school students and English teachers in Saudi Arabia. Thus, although the subjects or the study are all from Buraydah, Saudi Arabia, the targeted population is all third intermediate level students and their teachers in Saudi Arabia. This will help to establish an acceptable level of external validity, especially when the study is replicated in different parts of Saudi Arabia.

Results

Results of the Quantitative Analysis of the First Research Question

1. Is using the TBLT method for teaching English as a second language for male third-grade students in intermediate schools in Saudi Arabia more effective in the acquisition of the English language, in terms of students' achievement on reading comprehension, than using the traditional "prompting" method?

Pretest summary statistics. The average score on the pretest for all the 122 students who took the test was 19.885 with minimum and maximum scores of 11.00 and 30.00 respectively. The standard deviation for the pretest scores was 4.03. Table 4 presents summary statistics for the pretest broken down by the two groups of the treatment.

Table 4.
Pretest Summary Statistics

Treatment Groups	*n*	*M*	*SD*
TBLT Group	66	19.561	4.218
Control Group	56	20.268	3.802

An independent t-test was conducted to answer the question, "Is there a significant difference in the pretest score between the TBLT and the traditional teaching methods groups?" There was no statistically significant differences, (t (120) = -.965, p= .336) between students taught by the TBLT method and students taught by the traditional method of teaching on their pretest score suggesting that the two groups' initial proficiency of the English reading comprehension before administering the treatment is about the same.

Posttests summary statistics. There are five sets of posttests. Each set consists of a standardized test and a researcher-prepared assessment.

Table 5.
Standardized Posttests Summary Statistics

Treatment Groups		Posttest 1	Posttest 2	Posttest 3	Posttest 4	Posttest 5
TBLT Group (n = 57)	*M*	7.035	5.246	6.632	5.719	7.070
	SD	.906	1.675	1.046	1.998	1.226
Control Group (n = 47)	*M*	1.128	1.904	5.617	1.723	3.298
	SD	1.498	1.421	1.739	1.470	1.559

All five standardized posttests have a scale that ranges from zero to eight. The other five researcher-prepared assessment posttests are measured on a scale ranges from zero to four. Table 5 provides summary statistics for the five standardized posttests broken down by the two treatment groups.

Table 5 shows that there are pronounced differences between TBLT and control groups across all the five posttests measures, with the exception of the third posttest. The largest difference, ($M_{TBLT} - M_{Control} = 5.907$) between the two groups was on the first posttest. The smallest difference ($M_{TBLT} - M_{Control} = 1.015$) between the two groups was on the third posttest.

Summary statistics for the five researcher-prepared posttests broken down by the two treatment groups are presented in Table 4. Similar to standardized posttests there are sizable differences between the TBLT and control groups on the five researcher-prepared posttests with the exception of the third posttest. The largest difference ($M_{TBLT} - M_{Control} = 1.014$) between the two groups was on the fourth posttest. The smallest difference ($M_{TBLT} - M_{Control} = - 0.119$) between the two groups was on the third posttest.

Table 6.
Researcher-Prepared Posttests Summary Statistics

Treatment Group		Posttest 1	Posttest 2	Posttest 3	Posttest 4	Posttest 5
TBLT Group (n = 57)	*M*	2.228	2.667	2.360	3.237	3.239
	SD	1.161	.970	1.125	.808	.872
Control Group (n = 47)	*M*	1.745	1.692	2.479	2.223	2.117
	SD	1.117	1.337	1.402	1.250	1.134

Results for treatment effect. Treatment effect makes up the major part of the quantitative analysis of this study in addressing the first general question. The adopted Split-Plot design for this analysis allows for answering several sub-questions that collectively address the general research question.

Standardized posttests results. A mixed Split-Plot design with one between-groups (TBLT teaching method versus traditional teaching method) factor and one within-subjects (standardized posttest1 to posttest5) factor plus a pretest was adopted to answer the research question.

Pretest effect. A between-subjects test for the pretest (F (1, 101) = 25.260, p=.000) revealed that there is a significant effect of the pretest across the five posttests with a large effect size (*partial* η^2 = .200). From this test, we can infer that having the pretest in the model contributed significantly to the model and any derived conclusion from the treatment effect is adjusted for this significant contribution of the pretest. Further look at the within-subjects test for examining whether the effect of the pretest varies significantly across the five posttests, (F (3.411, 344.509) = 2.269, p=.072) revealed no statistical significant for this variation. This test infers that the pretest effect on the posttests does not vary significantly across the five posttests. Both of the between and within-subjects tests indicate that the pretest effect on the posttests is invariant across the five posttests.

Treatment effect. The between-subjects test for examining the treatment effect indicates that there is a significant treatment effect on the posttests scores (F (1,101) = 518.311, p= .000) with a relatively large effect size (η^2= .837). Students taught with TBLT method on average scored (M=6.373, SE=0.108) higher across the posttests than students taught with the traditional teaching method (M=2.694, SE= .119).

Looking at the treatment effect within the five standardized posttests (within-subjects effect) showed a significant interaction with the five posttests (F (3.411, 344.509) = 45.701, p= .000) with a large effect size (η^2= .312). Based on the estimated model Table 5 shows that the largest difference between TBLT (M = 7.050, SE = .160) and control (M = 1.110, SE = .176) groups occurred on the first standardized posttest. The smallest difference was found on the third posttest where TBLT students (M = 6.667, SE = .178) on average scored a bit higher than control group students (M = 5.574, SE = .196).

Table 7.
Standardized Posttests Estimated Means and Their Standard Errors

Treatment groups		Posttest 1	Posttest 2	Posttest 3	Posttest 4	Posttest 5
TBLT	*M*	7.050	5.285	6.667	5.774	7.091
(n = 57)	*SE*	.160	.198	.178	.220	.182
Control	*M*	1.110	1.856	5.574	1.657	3.372
(n = 47)	*SE*	.176	.218	.196	.242	.200

Note. TBLT= Task-Based Language Teaching

Figure 1 depicts the estimated means in Table 7. The graph on the next page shows that TBLT students, on average, performed better than control group students on all the five posttests with varying degrees.

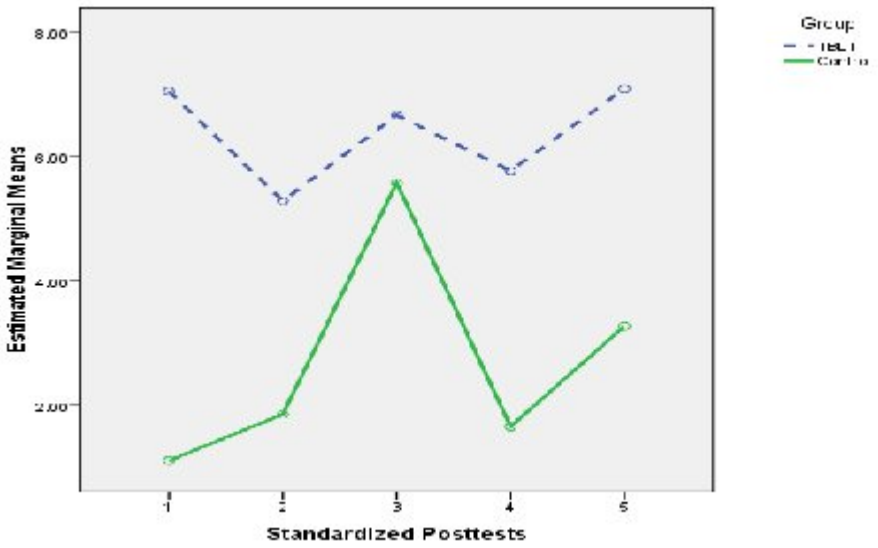

Figure 1. Estimated Means of Standardized Posttests for TBLT and Control Groups

Posttest effect. One aspect of the analysis is examining the pattern of the differences among the posttests regardless of the group (TBLT vs. control) membership. The within-subjects effect for testing the differences between the five posttests is statistically significant (F (3.411, 344.509)= 6.252, p= .000) indicating that, on average, students' posttests scores do differ significantly from one posttest to another with a medium effect size (*partial* η^2= .058). Table 6 presents the five standardized posttests estimated means, their standard errors of estimation, and the 95% confidence interval associated with each estimated mean.

Table 8
Standardized Posttests Estimated Means with Their Standard Errors and 95% Confidence Interval

Posttest	M	SE	(95% CI)
1	4.080	.118	(3.845, 4.315)
2	3.571	.147	(3.279, 3.863)
3	6.121	.132	(5.859, 6.382)
4	3.716	.163	(3.392, 4.039)
5	5.182	.135	(4.914, 5.449)

The 95% confidence intervals for the posttests' means show that all of the posttests means are significantly different from zero. Figure 6 displays the estimated means of the five posttests. Students' highest score was on the third posttest when compared to the remaining four posttests.

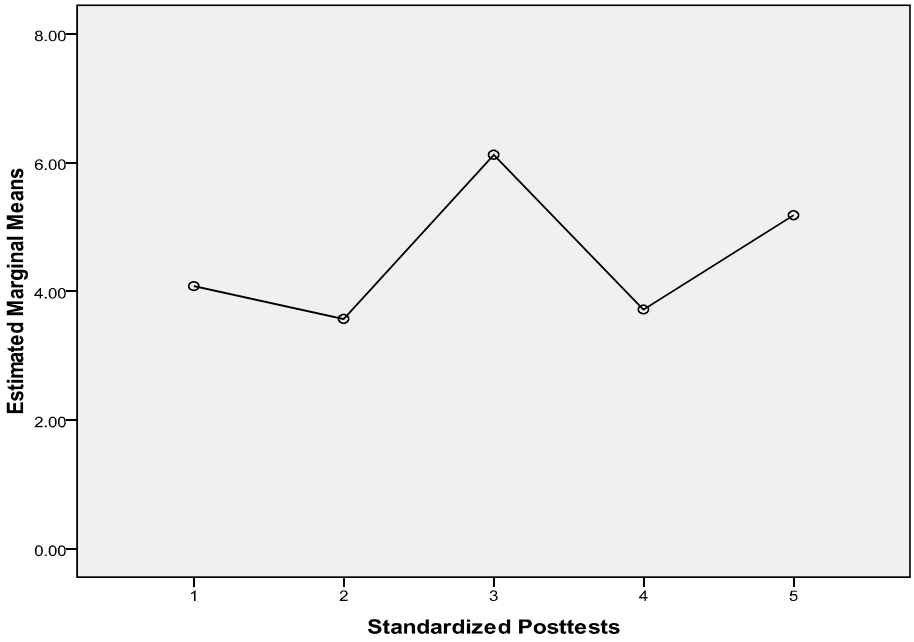

Figure 2. Estimated Means for the Five Standardized Posttests.

The fact that the five posttests were conducted over a period of 10 weeks with two weeks apart, allows for further investigation for the presence of possible significant trends in students' scores over time. There was a significant quadratic trend (F (1, 101) = 5.107, p = .026) and Order 4 trend (F (1, 101) = 19.264, p = .000). While both trends are feasibly possible to represent the fluctuation in the posttests means, Order 4 appears to be the representation of that fluctuation (i.e., note the p value). The posttest means tend to decline on the second posttest, incline sharply on the third, decline again on the fourth, then moderately incline on the fifth posttest giving us the significant Order 4 trend.

Bonferroni pairwise multiple comparisons of the posttests means gives another closer look at the posttests means' fluctuation. Table 7 presents the ten pairwise comparisons among the posttest scores with their statistical significance. Eight out of the ten pairwise comparisons were large enough to be statistically significant. The largest significant difference in posttests scores was between the second and the third posttests (-2.550, with p= .000). Out of the two insignificant pair wise comparisons, the smallest difference was between the second and the fourth posttests scores (-.145, with p= 1.000).

Table 9
Bonferroni Pairwise Comparisons Among Standardized Posttests

Posttest	2	3	4	5
1	.509*	-2.041***	.364	-1.102***
2		-2.550***	-.145	-1.611***
3			2.405***	.939***
4				-1.466***

Note. *p < .05, **p < .01, ***p < .001.

Researcher-Prepared Posttests Researcher

Similar analysis for standardized posttests is used, Split-Plot design, to analyze researcher-prepared posttests score. The analysis examines the effects of the three factors in the split-plot design, which are the between-groups (TBLT teaching method versus traditional teaching method) factor and the within-subjects (researcher-prepared posttest1 to posttest5) factor plus controlling for a pretest as a covariate in the model.

Pretest effect. A between-subjects test of the pretest revealed that there was a significant effect of the pretest across the five researcher-prepared posttests (F (1, 101) = 13.611, p= .000) and effect size (η^2 = .119) This test infers that having the pretest in the model contributed significantly to the model and any derived conclusion from the treatment effect on researcher-prepared posttests scores is adjusted for this significant contribution of the pretest.

A within-subjects test for examining if the effect of the pretest varies significantly within the five researcher-prepared posttests revealed no statistical significant for this variation (F (4, 404) = .106, p= .980) with a small effect size (η^2= .001). This test indicates that the pretest effect on the posttests does not vary significantly across the five posttests. Both of the between and within-subjects tests suggest that the pretest significant effect on the posttests is invariant across the five posttests. These findings are similar to those obtained when using standardized posttests scores. That is the pretest does have about the same significant effect on the five posttests.

Treatment effect. The between-subjects test for examining the treatment effect indicated that there is a significant treatment effect on the posttests scores (F (1,101) = 24.483, p= .000) with relatively small to moderate effect size (η^2=0.195). Students taught with TBLT method on average scored (M=2.768, SE= .101) higher across researcher-prepared posttests than students taught with the traditional teaching method (M=2.024, SE= .111).

Looking at the treatment effect within the five researcher-prepared posttests (within-subjects effect) revealed a significant interaction of the treatment with the five posttests (F (4, 404) = 9.061, p= .000) with a medium to a large effect size (η^2= .082).

Table 10.

Researcher-Prepared Posttests Estimated Means and Their Standard Errors

Treatment groups		Posttest 1	Posttest 2	Posttest 3	Posttest 4	Posttest 5
TBLT	*M*	2.247	2.691	2.382	3.260	3.261
(n = 57)	*SE*	.149	.148	.164	.132	.127
Control	*M*	1.722	1.662	2.452	2.195	2.088
(n = 47)	*SE*	.164	.163	.180	.145	.140

Note. TBLT= Task-Based Language Teaching

Figure 3 on the next page depicts the estimated means in Table 10.

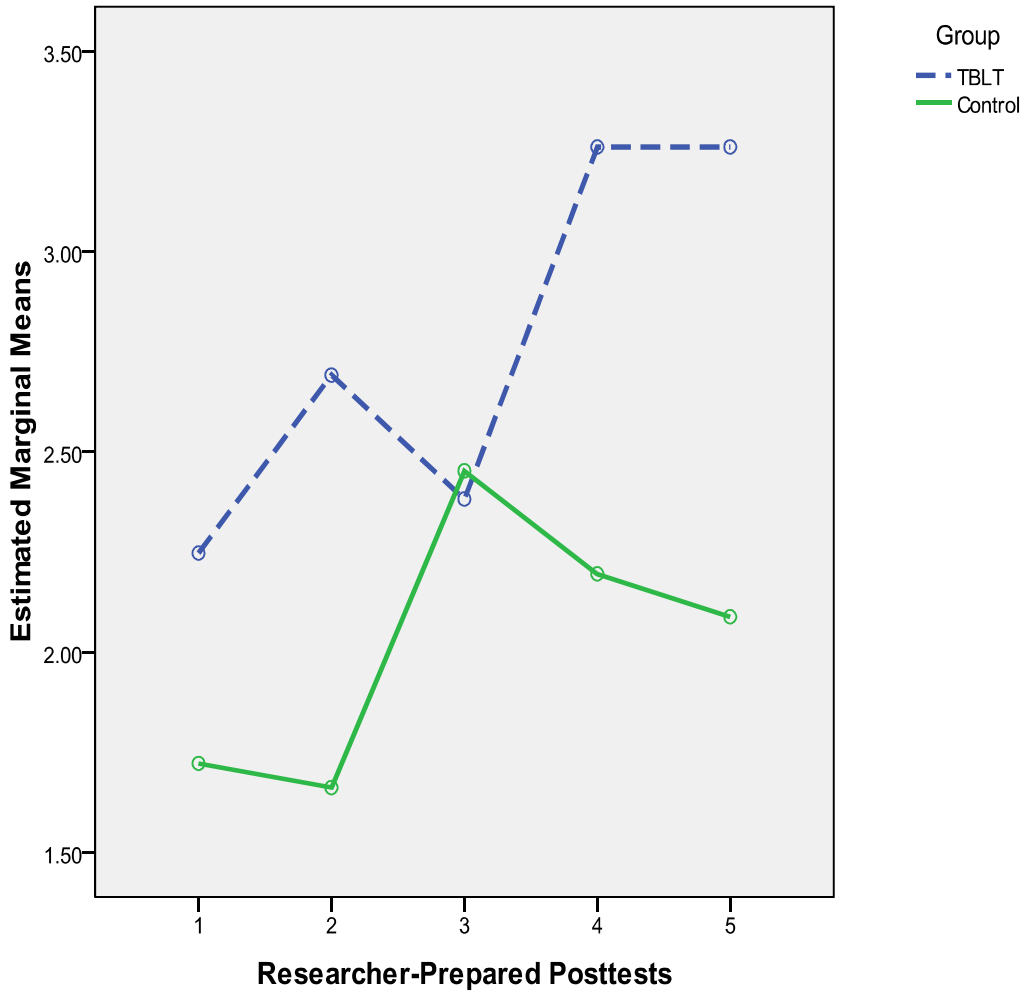

Figure 3. Estimated Means of Researcher-Prepared Posttests for TBLT and Control Groups

Based on the estimated model, Table 10 shows that the largest difference between TBLT (*M* = 3.261, *SE* = .127) and control (*M* = 2.088, *SE* = .140) groups occurred on the fifth researcher-prepared assessment. The smallest difference was found on the third posttest where TBLT students (*M* = 2.382, *SE* = .164) on average scored a bit lower than control group students (*M* = 2.452, *SE* = .180). The graph shows that TBLT students, on average, performed better than control group students did on four of the five posttests. Students from both groups scored about the same with slightly higher scores in favor of the control group on the third researcher-prepared posttest.

Posttest effect. As in the analysis of standardized posttests, the following shows examination of the pattern of the differences among researcher-prepared posttests regardless of the group (TBLT vs. control) membership. The within-subjects test for examining the differences between the five posttests is not statistically significant indicating that, on average, students' posttests scores did not differ significantly from one posttest to another (*F* (4, 404)= .341, *p*= .850) and a small effect size (η^2= .003). Table 9 presents the five researcher-prepared posttests estimated means, their standard errors of estimation and the 95% confidence interval associated with each estimated mean.

Table 11
Researcher-Prepared Posttests Estimated Means with Their Standard Errors and 95% Confidence Interval

Posttest	M	SE	(95% CI)
1	1.984	.111	(1.765, 2.204)
2	2.176	.110	(1.959, 2.394)
3	2.417	.122	(2.176, 2.658)
4	2.728	.098	(2.533, 2.922)
5	2.674	.094	(2.487, 2.862)

The 95% confidence intervals for the posttests mean show that all of the posttests means are significantly different from zero. Figure 8 displays the estimated means of the five posttests. Students' highest score was on the third posttest when compared to the remaining four posttests.

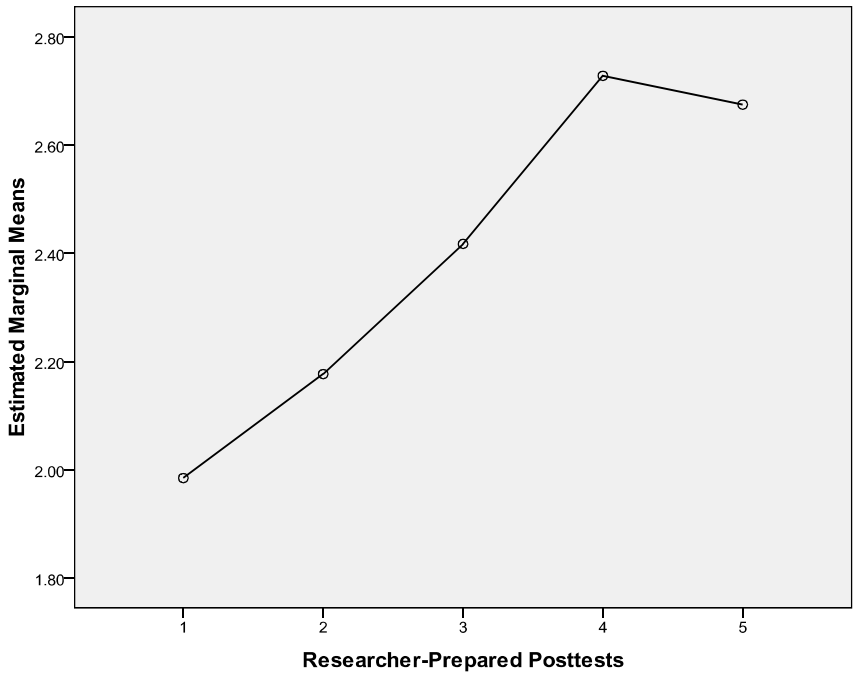

Figure 4. Estimated Means for the Five Researcher-Prepared Posttests

Table 12 presents the ten pairwise comparisons among researcher-prepared posttests scores with their statistical significance. Five out of the ten pairwise comparisons were large enough to be statistically significant. The largest significant difference in posttests scores was between the first and the fourth posttests (-.743, with p= .000). Out of the remaining five insignificant pairwise comparisons, the smallest difference was between the fourth and the fifth posttests scores (.053, with p= 1.000). Interestingly enough when the five significant contrasts depicted on the graph in Figure 8, it is obvious that there is a consistent gradual pattern of significant differences. Differences between the posttests start to be significant as we move from the first to the third posttests and gradually continue to be significant all the way to the fifth posttest. Difference between the first and the second posttests was not large enough to be statistically significant. Similar pattern inhibited in the second row of Table 10,

where the significant differences started in comparing the second posttest with the fourth and the fifth posttests. Comparison between the second and the third posttests were deemed not to be significant. Such pattern does not exist when comparing the third posttest to the fourth or the fifth posttests as the means in these comparisons are from posttests either next to each other or very close. This is also true for the last comparison between the fourth and the fifth posttests. Furthermore, all differences have negative values indicating that there is a gradual improvement in students' performances over time except between the fourth and the fifth posttests where that difference was very small in magnitude yet positive value. This difference is not significant and could be an artifact of a random error in the sample.

Table 12.
Bonferroni Pair Wise Comparisons Among Researcher-Prepared Posttests

Posttest	2	3	4	5
1	-.192	-.432*	-.743***	-.690***
2		-.240	-.551***	-.498***
3			-.311	-.258
4				.053

Note. $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

Results of the Qualitative Analysis of the Second Research Question

2. What insights and issues can be gained about implementing TBLT in this research setting?

Observational visits to the control group. Findings interpreted out of observational data associated with the control group have shown trends, attitudes, advantages, and disadvantages. One of these trends is that teaching via the traditional method is *monotonous.* In other words, the way the lessons were introduced, run, and assessed, students and teacher's attitudes, difficulties, advantages, and disadvantages were pretty much similar across most of the observational visits. The monotonous nature of the traditional teaching method yielded almost similar observational notes for the students and teacher's attitudes, difficulties, and advantages and disadvantages.

Students' attitude lacked interest in the reading lessons. This interpretation is obtained from a number of responses and actions done by the students across the several visits. For examples, eyes wandering in the ceiling and sometimes yawning of students were pretty much common across most of the observational visits. In the third week, one student said quietly to his neighbor, "do we have to study this?" showing little care to be overheard by the investigator who was sitting next to them.

Teacher's attitudes, on the other hand, provided another evidence of inadequacy of the traditional teaching method. The teacher was tensed with the situation when students were not paying the expected attention across most of the observed reading lessons and, hence, repeatedly asked students, in a tune that showed inconvenience, to follow with him and concentrate at the reading passage.

The advantages of the traditional teaching method seemed to be far less than the observed disadvantages. One advantage of the traditional teaching method was that it was easy for the teacher to teach and enabled him to have control over class the entire duration of

the lesson. Another controversial advantage was that students were quiet across most of the lessons observed.

Disadvantages, on the other hand, could be seen from three main perspectives. The first one was that the traditional teaching method was highly *teacher-centered*. In other words, it was the teacher who did most of the work in the reading lesson. The teacher would read the passage, explain the reading passage, assign two or three students to read, give instructions to students, and read exercises and ask students to do them. The teacher alone used about 70 to 80% of time of the duration of the reading lesson. The second perspective was that the students were bored with the English language reading class. Students used from 70 to 80% of time listening to their teacher while speaking. Students did not have any types of activities to do during the reading lesson except the one they do individually towards the end of the lesson. The third perspective of disadvantages was that the traditional teaching method heavily relied on prompting practices. In other words, instruction and explanation were always orally by the teacher. Students also did several drills to memorize the correct pronunciations of some English words.

Observational visits to the TBLT group. Findings interpreted out of observational data associated with the TBLT group have shown trends, attitudes, advantages, and disadvantages**.** One of the trends was that teaching via the TBLT method took longer time to describe. In other words, written notes about the teaching and learning situations while implementing TBLT had more descriptive details. The reason behind that is that the nature of the TBLT method consists of various elements that lead to more actions to take place in classroom from all parties involved in the lesson being taught. In essence, there were more things that took place while running the lesson and, hence, needed more words to describe them.

Another interpreted trend about the implementation of TBLT was consistency. In other words, the *skeleton* of the reading lesson taught via TBLT consisted of three main stages. The first one, *pre-task,* aimed at engaging students into the main goal of the lesson. The second stage, *running the task,* described students while they were actually doing what they were intended to do. In the third stage, *task completion,* students provided their teacher with a product for the purposes of assessing to what extent students had achieved the main goal of the lesson.

Having said that the observational data suggested consistency following the three stages of the TBLT method, teaching and learning were also characterized by having a great deal of variety. In other words, various activities took place during the three fixed stages of the TBLT method. For example, in a reading lesson in the second week about *Calvin Hutt's Career Life,* students in the *pre-task* stage provided their classmates with lists of video games they were playing at home and read a passage about *Calvin Hutt's Career Life* in the *running task* stage. Students in the *task completion* stage imagined they were participating in a live competition show to answer a question asked by the interviewer where they told the audience (their teacher and other groups of students) as much details as they could about *Calvin Hutt's Career Life*.

The most prevailing trend across most of the nine observational data categories was that learning via the TBLT method was *learner-centered.* Learner-centered meant here that that the students were the central focus of instruction and students participated in creating their learning situations. To clarify this notion, a careful investigation is bestowed to the nine observational data categories. Students were described or mentioned by the observer almost

in every cell across all columns unlike the teacher whom the observer mentioned fewer times and described in roles of being a facilitator rather than a source of instruction. In other words, students were *active learners* (i.e., they were discussing, negotiating, reading, and displaying their understanding of what they had been learning). This meant the learning situation via the TBLT method revolved around the learners.

Students had realized in the first week of the study the difference occurred in the way they were taught and which appeared in one student's comment to his group, "we are studying differently." Studying via TBLT or "studying differently" had positively enhanced students' verbal responses towards the learning situation, and which was revealed in multiple occasions across the following weeks of the study. For instances, students tended to organize themselves at the beginning of each lesson, join their groups, and show readiness to start the reading lessons without much efforts or further notices from the treatment teacher (researcher). Also, the students always showed engagement in group works and enthusiastically shared their responses with their classmates.

Careful analysis of the two categories of observational data related to the advantages and disadvantages of the implementation of the TBLT method showed that the advantages and disadvantages went along with or supported by the interpreted trends earlier. Examples of advantages related to learners included; that students were very active in terms of asking questions and sharing responses, negotiation of meaning was always present among groups of students while reading passages, focus was on students since they tended to speak far more than their teacher, presence of peer or collegial learning as students learned more details about the reading passage from shared responses by groups of students, and students' comprehension of meaning was always the ultimate aim targeted by the practices involved in the TBLT lesson. Interpreted advantages related to the work of the teacher were much less than the observed ones about the students. The reason was that the teacher was not the central focus or the main source of information and, hence, focus was more on the students who were making action. Among the advantages that described the work of the teacher included that he was modeling the role of a facilitator as he was passing among groups providing them with guidance, monitoring group works, relaxed, and frequently used his sense of humor.

Disadvantages were minimal and related to the work of the teacher in the classroom rather than that of students. The most prevailing disadvantage about the implementation of the TBLT method was that it was demanding on the teacher and required mental and physical attendance by the teacher. At last but not least, implementing TBLT required more time and, hence, any unintended loss of time might easily lead to failure to achieve the main goal of the lesson as planned. At last, teaching via the TBLT method was not easy work for the teacher and required certain skills and background about the TBLT method before implementing it on the classroom, and which the treatment teacher had while he was teaching.

Researcher Log

The analysis of data collected via *researcher log* showed that they revolved around four categories. Not surprisingly, the two most prevailing categories were about the students' roles and attitudes in the classroom. The collected data under those two categories support the observational data interpreted under Table 12 and which adds further strength to the findings. The third category was very much related to running TBLT as a method of teaching with emphasis on reading comprehension in an English as a second language classroom in this research setting. The fourth category of data was related to the role and impression of the

teacher (researcher) in the classroom while he was teaching via the TBLT method in this research setting. *Figure 9* shows the numbers of counted *key words* that describe each category.[1]

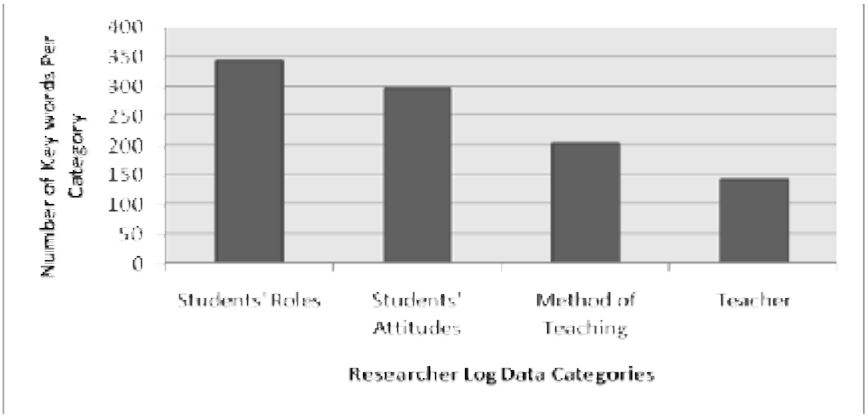

Figure 5. Number of Counted Key Words under the Four Categories

The reasons that made students' roles and attitudes be the two most prevailing categories in a reading lesson taught via the TBLT method could be linked to the reality that they had played a central role or were the action makers during the flow of the reading lesson. Data out of *researcher log* repeatedly described the roles of students they had been playing in the classroom. The three most frequent roles included reading, discussing, and sharing. Collected data also tended to report what students had been doing in the classroom. Students started every reading lesson with a group discussion, *the pre-task stage*, involving an activity that imitated students' daily lives and which helped engage the students in the intended reading content. Every group of students extended its work after the engaging activity to read the reading passage, group members discussed their understanding of what they had read, and formulated an agreed-upon response to share with other groups.

The second half of the prevailing data was a description of students' attitudes towards the reading lesson. The two most common words used to describe the students' attitudes were *enthusiasm* and *involvement*. Enthusiasm and involvement in this setting referred to the manner in which the students were performing the aforementioned roles (reading, discussing, and sharing). Two examples derived from the raw data could give a picture of enthusiasm and involvement of students in the reading lesson. In the third week, two groups of students had an argument about who should have the turn first to start sharing their responses with other groups. The group that started first usually had the opportunity to share another time as long as the time allotted for group sharing was not consumed. The second example was about a student who actually broke the boundaries of group work in the fifth week. When the turn was for his group to share with a response, that student enthusiastically stood up and orally narrated the whole of the reading passage consuming more than the time allotted for his group. In the meantime, the teacher (researcher) tried to politely give the chance to another group but the student would not stop and continued all the way to the end of his long response.

[1] Figure 9 includes counted numbers of key words under each category and excludes neutral words such as articles, prepositions, and verbs to be.

Students' attitude towards the reading lesson taught via the TBLT method was positive. Interpreted data showed that they even loved and enjoyed what they were doing in the reading lesson. Beside the never observed complain or lack of interests tokens that universally accompany any an undesired class by students at the age of the students participating in the study, the treatment teacher (researcher) considered an incident that had happened in the fourth week as an evidence or at least an indicator that the students loved the reading lesson taught via the TBLT method.

It was Wednesday (the last day of school week days in Saudi Arabia) when the teacher (researcher), as usual after teaching students the reading class, headed towards the teachers' office room. The counselor stopped the treatment teacher and asked him if he gave his instructions to five of the students not to participate in a tour outside the school hosted by an outsider organization. The treatment teacher told the counselor that he had not given any instructions in this regards. In the middle of that conversation, the treatment teacher was shocked out of surprise and wondered about the reason that prevented the five students from going on the tour as he knew that every student wished to participate in similar tours. The treatment teacher asked the counselor about the reason that made the students opted not participate in the tour. The counselor replied that the students said that they had had an English reading class and they did not want go on the tour. That incident was complemented by a phone call after the end of the study by the original teacher to the treatment teacher (researcher) stating that some students asked him to teach them the way the treatment teacher (researcher) was teaching them.

The third category was related to the application of the TBLT method in this research setting. Interpreted data out of *researcher log* suggested some difficulties that the treatment teacher had faced when implementing the TBLT method. One of those difficulties was that the students did not know the meaning or not used to group work. At the beginning, students were sitting in groups but working individually which made the treatment teacher correct that at once explaining duties and expectations out of group work. Another difficulty, which might be a consequence of the first one, was the existence of minimal side talks (not related to the lesson at hand) among some students at the beginning of the study. However, as the study progressed and students understood the meaning of group work, those minimal side talks started to vanish. The last difficulty was related to the challenge of time. Time was congesting and reading lessons taught in this study tended to finish exactly by the end of the allotted duration of time and sometime a minute or two minutes were to be borrowed from the breaks following the lessons. That warned that any unintended loss of time might severely prevent students from achieving the goal of the lesson as planned.

The fourth category was related to the teacher's (researcher) role and impression in this research setting. Interpreted data out of *researcher log* showed that the treatment teacher (researcher) had described what he was doing in every class he had taught. The way he was teaching was consistent across all lessons and strictly followed the principles of the TBLT method he was implementing including the three stages of a TBLT lesson (*pre-task, running the task,* and *task completion*). The treatment teacher precisely followed the lesson plans he designed for every reading lesson class. To engage students in the main task of the lesson, those lessons tended to start with group activities that were derived from students' daily lives while ensuring the achievement of the main goal of the lesson (*running the task*) was through a *retelling* activity that too imitated students' real lives. The teacher's (researcher) impression showed always satisfaction about the way he taught and the way students were

working in the classroom. However, a couple of times the TBLT teacher mentioned that he was exhausted and that might be linked to the continuous physical motion the treatment teacher was doing while passing among groups and paying attention to groups' discussions as well as participating with them.

Discussion

Both the quantitative and qualitative analyses provided detailed description of the analysis and the findings of the study. While each analysis revealed specific findings that were related to the nature of the data collection tools used to answer one of the two research questions, this discussion attempts to combine findings of both quantitatively and qualitatively collected data to provide a full or complete picture about the findings of the study.

The overall of the statistical analyses of the quantitatively collected data provided valuable findings to answer the first research question. The major finding that explicitly answered this question was: yes, the application of the TBLT method for teaching English as a second language for male, third-grade students in intermediate schools in Saudi Arabia was more effective in the acquisition of the English language, in terms of students' achievement on reading comprehension, than using the traditional "prompting" method. In other words, the application of TBLT method in this research setting helped increase the students' achievement scores in reading comprehension. That major finding was statistically reported by two sources of data (researcher-prepared assessments and standardized text-established tests), and which even increased the validity of the findings. The pretest results also showed that students were equal across the control and treatment groups eliminating the possibility for effect of initial level of the English language reading comprehension before the application of the TBLT method. The average scores of both types of posttests (researcher-prepared assessment and standardized text-established tests) of the control and treatment groups were highly significant in favor of the treatment group.

Qualitatively collected data on the other hand greatly helped describe and explain the surroundings of the application of the TBLT method in this research setting. Because neither group knew their group identification (treatment or control) nor knew the way they were going to be taught before the beginning of the study, this study assumes that students in both groups have a very low level of possibility to form a prejudice that might interact with their attitudes towards the learning situations. Hence, the interpreted qualitative data showed that teaching via the TBLT method in this research setting helped students develop desired attitudes towards the learning situations, unlike the traditional teaching method that showed that students had developed undesired attitudes towards the learning situations as elaborately explained under the analyses of the qualitative data. Another vital finding interpreted from the qualitative data was that teaching via the TBLT method required both of the students and their teacher to play roles or involve in practices that went along with the practices of the constructivist learning theory, unlike the traditional teaching method which involved practices and roles of students and their teacher that went along with the behaviorist learning theory.

Interpreted quantitative and qualitative data when combined showed that they had provided support and evidences for the findings suggested by each set of data. In other words, qualitative findings that suggested that the TBLT method had helped the students in the treatment group develop desired attitudes towards the learning situations were supported by the quantitative findings that showed that the TBLT method had also helped students

increase their achievement scores in reading comprehension of the English language. Also, the quantitative findings that showed that the traditional teaching method did not help students in the control group increase their achievement scores in reading comprehension as compared to that of the TBLT method were supported by the qualitative findings that showed that the traditional teaching method also did not help the students develop desired attitudes towards the learning situations as that of the TBLT method.

Conclusion

The study has strived to find out whether or not the TBLT method can help the students better acquire the English language through increasing their achievement scores on reading comprehension and also seek for insights or issues that can be gained about implementing the TBLT method in this research setting. Literature reviewed has shown that the TBLT method is theoretically framed by the constructivist learning theory.

The study has examined the effect of TBLT on reading comprehension in two intermediate schools in Saudi Arabia through a time frame of ten weeks. The treatment group is compared to the control group on the outcome after controlling for the students' pre-existing knowledge of the English language as a covariate.

Findings out of the pre-test have shown that students in both of the treatment and control groups are equal in terms of their prior knowledge of reading comprehension of the English language. Findings out of the posttests have shown that students in the treatment group have scored significantly higher than those in the control group. Findings out of the qualitative data have shown that the TBLT method has helped students develop desired attitudes towards the learning situations and has involved practices and roles of students and their teacher that go along with the constructivist learning theory. Qualitative findings have also shown that the traditional teaching method has not helped students develop desired attitudes towards the learning situations and involved practices and roles of students and their teacher that went along with the behaviorist learning theory. The findings have provided support and evidences for the findings presented by each set of data.

References

Albert, A. & Kormos, J. (2004). Creativity and narrative task: An exploratory study. Language Learning. 54(2), 270-310. doi: 10.1111/j.1467-9922.2004.00256.x

Applegate, M. D., Quinn, K. B., & Applegate, A. J. (2008). The critical reading inventory: Assessing students' reading and thinking (2nd edition). Upper Saddle River, NJ: Pearson

Breen, M. (1989). The evaluation cycle for language learning tasks. Cambridge, UK: Cambridge University Press

Brown, H. D. (2007). Teaching by principles: an interactive approach to language pedagogy. New York, NY: Pearson

Burt, M., Dulay, H., & Hernandez, E. (1973). Bilingual syntax measure. New York: Harcout Brace Jovanovich

Bygate, M., Skehan, P., & Swain, M. (Eds.). (2001) Researching pedagogic tasks, second language learning, teaching and testing. Harlow, UK: Pearson

Candlin, C. N. (2001). Afterword: Taking the curriculum to task. In M. Bygate, P. Skehan, & M. Swain (Eds.), Researching pedagogic tasks: Second language learning, teaching, and testing (pp. 229–243). Harlow, UK: Pearson

Carless, D. (2008). Student use of the mother tongue in the task-based classroom. ELT Journal, 62(4), 331–338. doi:10.1093/elt/ccm090

Crookes, G. (1986). Task classification: a cross-disciplinary review. Honolulu: University of Hawaii

Cummins, J. (1983). Language proficiency and academic achievement. In J. Oller, (Ed.), Issues in language testing research (108-130). Rowley, Mass: Newbury House

De Bot, K. (2001). Interaction in the classroom. TESOL Quarterly, 35, 602–603. Retrieved from http://www.jstor.org/stable/3588434

Ellis, R. (2003). Task-based language learning and teaching. New York, NY: Oxford University Press

Fraenkel J. R., Wallen N. E., & Hyun, H. H. (2012). How to design and evaluate research in education (8th ed.). Boston: McGraw-Hill

Hackuta, K. (1976). Becoming bilingual: a case study of a Japanese child learning English. Language learning. 26, 321-351

Izadpanah, S. (2010). A study on task-based language teaching: From theory to practice. US-China Foreign Language, 8(3), 47–56. Retrieved from http://web.ebscohost.com/ehost/pdfviewer/pdfviewer?vid=3&hid=13&sid =63d2f07d-7805-4ee5-a5ac-85b487537f24%40sessionmgr10

Jarvis, P. (2006). Towards a comprehensive theory of human learning. New York, NY: Routledge

Jarvis, P. & Parker, S. (2005). Human learning: A holistic perspective. London: Routledge Falmer

Jarvis, P. (2009). Learning to be a person in society: Learning to be me. In K. Illeris, (Eds.), Contemporary theories for learning: learning theorists in their own words (pp. 21-34). London: Routledge

Kim, Y. (2008). The role of task-induced involvement and learner proficiency in L2 vocabulary acquisition. Language Learning, 58(2), 285–325. doi: 10.1111/j.1467- 9922.2008.00442.x

Krashen, S. (1994). The input hypothesis and its rivals' in N. Ellis (Ed.), implicit and explicit learning of languages (pp. 45-77). London: Academic Press

Kumaravadivelu, B. (2006). TESOL methods: changing tracks, challenging trends. *TESOL Quarterly, 40*, 59–81. Retrieved from http://www.jstor.org/stable/40264511

Lantolf, J. (2000). Sociocultural theory and second language learning. New York, NY: Oxford University Press

Lee, J. (2000). Tasks and communicating in language classroom. Boston, MA: McGraw-Hill

Long, M. (1996). The role of the linguistic environment in second language acquisition. In W. Ritchie, & T. Bhatia, (Eds.), Handbook of second language research (pp. 413-468). San Diego: Academic Press

Long, M. (1985). A role for instruction in second language acquisition: task-based language teaching. UK: Hyltenstam, M. Pienemann

Maroun, N. & Samman, H. (2008). How to succeed at education reform: the case for Saudi Arabia and the broader GCC region. Ideation Center, Middle East: Booz Allen Hamilton

McNiff, J., Lomax, P., & Whitehead, J. (1996). You and your action research project. New York: Routledge

Mohamed, N. (2004). Consciousness-raising tasks: A learner perspective. ELT Journal, 58(3), 228–237. Retrieved from http://journals.ohiolink.edu/ejc/pdf.cgi/Mohamed_Naashia.pdf?issn=09510893&issue=v58i0003&article=228_ctalp

Murphy, J. (2003). Task-based learning: The interaction between tasks and learners. ELT Journal, 57(4), 352–360. Retrieved from http://journals.ohiolink.edu/ejc/pdf.cgi/Murphy_Jacky.pdf?issn=09510893&issue=v57i0004&article=352_tltibtal

Nunan, D. (1989). Designing tasks for the communicative classroom. Cambridge, UK: Cambridge University Press.

Piaget, J. (1970). Science of education and the psychology of the child. New York: Viking

Plews, J. L. & Zhao, K. (2010). Tinkering with tasks knows no bounds: ESL teachers' adaptation of task-based language teaching. TESL Canada Journal, 28(1), 41–95. Retrieved from http://journals.sfu.ca/tesl/index.php/tesl/article/view/1059/878

Poehner, M. (2009). Group Dynamic Assessment: Mediation for the L2 Classroom. TESOL Quarterly, 43(3), 471-491. doi: 10.1002/j.1545-7249.2009.tb00245.x

Prabhu, N. S. (1987). Second Language Pedagogy. New York, NY: Oxford University Press

Richards, J., Platt, J., & Weber, H. (1985). Longman Dictionary of Applied Linguistics. London, UK: Longman

Rivers, D. (2010). An exploration of on-task language policy and student satisfaction. ELT Journal, 64(3), 261–271. doi:10.1093/elt/ccp05 Rivers, D. (2010). An exploration of on-task language policy and student satisfaction. ELT Journal, 64(3), 261–271. doi:10.1093/elt/ccp058

Robinson, P. (2001). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. Applied Linguistics. 22(1), 27–57. Retrieved from http://journals.ohiolink.edu/ejc/pdf.cgi/Robinson_P.pdf?issn=01426001&issue=v22i0001&article=27_tctdateiiacf

Samuda, V. & Bygate, M. (2008). Tasks in second language learning. Basingstoke, UK: Palgrave Macmillan

Schneider, D. (2005). Qualitative data acquisition methods: interviews and observation. Research Design for Educational Technologists. 1, 1-16. Retrieved from http://tecfa.unige.ch/guides/methodo/edu-tech/slides/data-quali.pdf

Schram, T. (2006). Conceptualizing and proposing qualitative research. New Jersey: Upper Saddle River

Skehan, P. (1998). A cognitive approach to language learning. New York, NY: Oxford University Press

Stevens, F. (1983). Activities to promote learning and communication in the second language classroom. TESOL Quarterly, 71, 259–272. Retrieved from http://www.jstor.org/stable/3586653

Swain, M. & Lapkin, S. (2000). Task-based second language learning: The uses of the first language. Language Teaching Research, 4(3), 251-274. Retrieved from http://web.ebscohost.com/ehost/pdfviewer/pdfviewer?vid=3&hid=13&sid=980d3cbc-87fc-4bbd-bcc2-9f182814363a%40sessionmgr12

Swain, M. & Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. The Modern Language Journal, 82(3) 320–337. Retrieved from http://www.jstor.org/pss/329959

Swan, M. (2005). Legislation by hypothesis: The case of task-based instruction. Applied Linguistics, 26(3), 376–401. doi:10.1093/applin/ami013

Vygotsky, L. (1978). Mind in society: the development of higher psychological processes. Cambridge. Harvard University Press

Wiersma, W. & Jurs, S. G. (2009). Research methods in education: An introduction (9th ed.). New York, NY: Library of Congress