

11-1-2018

Task-based Language Teaching versus Present, Practice, Produce: Efficacy in Language Learning and Assessment

Majeed Noroozi

Florida International University, mnaro002@fiu.edu

Follow this and additional works at: <https://digitalcommons.fiu.edu/etd>



Part of the [Curriculum and Instruction Commons](#)

Recommended Citation

Noroozi, Majeed, "Task-based Language Teaching versus Present, Practice, Produce: Efficacy in Language Learning and Assessment" (2018). *FIU Electronic Theses and Dissertations*. 3874.
<https://digitalcommons.fiu.edu/etd/3874>

This work is brought to you for free and open access by the University Graduate School at FIU Digital Commons. It has been accepted for inclusion in FIU Electronic Theses and Dissertations by an authorized administrator of FIU Digital Commons. For more information, please contact dcc@fiu.edu.

FLORIDA INTERNATIONAL UNIVERSITY

Miami, Florida

TASK-BASED LANGUAGE TEACHING VERSUS PRESENT, PRACTICE,
PRODUCE: EFFICACY IN LANGUAGE LEARNING AND ASSESSMENT

A dissertation submitted in partial fulfillment of the
requirements for the degree of

DOCTOR OF PHILOSOPHY

in

CURRICULUM AND INSTRUCTION

by

Majeed Noroozi

2018

To: Dean Michael R. Heithaus
College of Arts, Science, and Education

This dissertation, written by Majeed Noroozi, and entitled Task-Based Language Teaching versus Present, Practice, Produce: Efficacy in Language Learning and Assessment, having been approved in respect to style and intellectual content, is referred to you for judgment.

We have read this dissertation and recommend that it be approved.

Teresa Lucas

Haiying Long

Eric Dwyer, Co-Major Professor

Melissa Baralt, Co-Major Professor

Date of Defense: November 1, 2018

The dissertation of Majeed Noroozi is approved.

Dean Michael R. Heithaus
College of Arts, Science, and Education

Andrés G. Gil
Vice President for Research and Economic Development
and Dean of the University Graduate School

Florida International University, 2018

© Copyright 2018 by Majeed Noroozi

All Rights Reserved

DEDICATION

Dedicated to my Parents

ACKNOWLEDGMENTS

I would like to thank Florida International University, School of Education for giving me the unique opportunity to afford and do my doctoral studies in Curriculum and Instruction. I am grateful for Dr. Melissa Baralt's continuous help and support throughout the painstaking process of conducting this dissertation. My special thanks go to Dr. Dwyer who has always been there for me and without whose support I could not have finished this study. In addition, I would like to thank my dissertation committee members, Dr. Long and Dr. Lucas for their invaluable advice and consultation towards the completion of this study.

ABSTRACT OF THE DISSERTATION

TASK-BASED LANGUAGE TEACHING VERSUS PRESENT, PRACTICE,
PRODUCE: EFFICACY IN LANGUAGE LEARNING AND ASSESSMENT

by

Majeed Noroozi

Florida International University, 2018

Miami, Florida

Professor Melissa Baralt, Co-Major Professor

Professor Eric Dwyer, Co-Major Professor

Long (2015) defines Task-based Language Teaching (TBLT) as “an approach to course design, implementation, and evaluation intended to meet the communicative needs of diverse groups of learners” (p. 5). Task-based Language Teaching has been introduced and developed by second language acquisition researchers as well as language educators in response to the teacher dominated and focus-on-formS methods of language teaching in classrooms such as the approach of Present, Practice, Produce (PPP) (Van den Branden, 2006). The present study aimed to build upon the previous literature on the possible differential effects of the PPP approach and TBLT on students’ language learning (e.g., De la Fuente, 2006;; Lai, Zhao, & Wang, 2011; Li, Ellis & Zhu, 2016; Gonzalez-Lloret & Nielson, 2015; Shintani, 2011, 2013) which have examined the differential effects of these two language methodologies on learners’ language learning. The present study aimed to address the methodological drawbacks of the Li et al. (2016) study by including Task-based Language Assessment (TBLA) in its methodology alongside the GJT and the EIT so as to obtain a more comprehensive picture of the comparison of PPP and TBLT.

Thirty-four participants from three English as a Foreign Language (EFL) classes at the lower intermediate level of proficiency participated in this study, which took place at the Parsian Language Institute located in the city of Ghaemshahr in Iran. The three classes were randomly assigned to three groups of TBLT, PPP, and Control. Learning was measured with the same types of tests as the Li et al. (2016) study, i.e., a GJT and an Elicited imitation test; however, a Task Assessment was added. Participants were administered the pre-assessments, then participated in the TBLT, PPP and Control group treatments, respectively, and finally performed the post-assessments. A Wilcoxon Signed Ranked Test revealed that the performance of TBLT and PPP on the GJT and the EIT significantly improved from pre-assessment to post-assessment, while the Control group did not show any significant improvements on any of the tests. As for the task assessment, results showed that only the TBLT group made significant improvements on their post-assessment, while the PPP and Control group did not statistically improve

TABLE OF CONTENTS

CHAPTER	PAGE
1. INTRODUCTION: PPP AND TBLT	1
1.1. Background of the Issue.....	1
1.1.1. TSLT vs. TBLT	3
1.2. Theoretical Rationale for TBLT	6
1.2.1. Interaction, Input, and Output Hypothesis	9
1.2.2. Focus on Form.....	10
1.2.3. Needs Analysis	12
1.3. Criticisms of TBLT.....	13
1.4. Task-based Language Assessment (TBLA).....	19
1.5. The Difference between TBLT and TBLA.....	22
1.6. Advantages of Task-based Language Assessment	23
1.7. The PPP Approach.....	23
1.8. Explicit/Declarative vs. Implicit/Automated knowledge.....	26
1.9. Definition of Key Terms.....	31
2. TASK-BASED LANGUAGE ASSESSMENT.....	33
2.1. Paradigms of Language Assessment.....	33
2.2. Performance Assessment	35
2.3. Task-based Performance-referenced Tests	37
2.3.1. Assessment in Task-based Performance-referenced Tests.....	38
2.4. Task-based System-referenced Tests.....	40
2.5. Direct vs. Indirect Tests	41
2.6. Measurement in Task-based Language Assessment.....	44
2.7. Issues in Task-based Language Assessment.....	45
3. REVIEW OF THE LITERATURE	49
3.1. TBLT vs. PPP	49
3.2. The Role of the Teacher.....	74
3.3. Statement of the Problem.....	75
3.3.1. Issues in the PPP Approach.....	75
3.3.2. The Importance of TBLA.....	76
3.3.3. Teacher's Role in TBLT.....	78
3.4. Purpose of the Study	80
3.5. Significance of the Study.....	82
3.6. Research Questions	83
3.7. Limitations of the Study.....	84
4. METHODS	86
4.1. Participants.....	86
4.2. Teachers	87
4.3. Instrumentation	87
4.3.1. Pre-assessment Instruments.....	87

4.3.1.1. The Grammaticality Judgement Test	88
4.3.1.2. The Elicited Imitation Test	89
4.3.1.3. Task Assessment	90
4.3.1.4. Coding	90
4.3.3. Treatment Instruments	91
4.3.3.1. The TBLT Tasks	91
4.3.3.2. The PPP Activities	92
4.3.3.3. The Control Group Activities	94
4.4. The Design of the Study	94
4.5. Procedure	96
4.5.1. Pilot Studies	96
4.5.2. Pre-assessment	98
4.5.3. The PPP Treatment	100
4.5.4. The TBLT Treatment	103
4.5.5. The Control Group Treatment	106
4.5.6. Post-assessment	107
5. RESULTS	110
5.1. Test of Normality	111
5.2. Pre-assessment Analysis	111
5.3. Non-parametric Analysis	112
5.3.1. The Grammaticality Judgement Test	112
5.3.2. The Elicited Imitation Test	115
5.3.3. Task Assessment	118
5.4. Post-assessment Analysis	120
6. DISCUSSION	122
6.1. The Grammaticality Judgement Test and Explicit Knowledge	123
6.2. The Elicited Imitation Test and Implicit Knowledge	125
6.3. Task Assessment and Task Performance	129
6.4. Task-Based Language Assessment	132
6.5. Comparison with Li et al. (2016) Study	135
6.6. Task-Based Context Knowledge and Discourse Choice	137
7. CONCLUSION	140
7.1. The GJT, the EIT, and Task Assessment	140
7.2. Task-based Language Assessment	146
7.3. Teachers and Classroom Research	147
7.4. Reflection	148
7.5. Limitations	149
7.6. Pedagogical Implications	151
7.7. Theoretical Implications	152
7.8. Suggestions for Future Research	153
LIST OF REFERENCES	155

APPENDICES	159
VITA.....	182

LIST OF TABLES

TABLE	PAGE
1. Modes of Language Assessment	43
2. Summary of Studies Comparing the PPP Approach with TBLT	70
3. Shapiro-Wilk Test of Normality.....	111
4. Kruskal-Wallis Test for Pre-Assessment Analysis.....	112
5. Means and Standard Deviations for the Grammaticality Judgement Test	113
6. Wilcoxon Signed Rank Test for the Grammaticality Judgement Test	115
7. Means and Standard Deviations for the Elicited Imitation Test.....	116
8. Wilcoxon Signed Rank Test for the Elicited Imitation Test	117
9. Means and Standard Deviations for Task Assessment.....	118
10. Wilcoxon Signed Rank Test for Task Assessment.....	119
11. Kruskal-Wallis Test for Post-Assessment Analysis	121

LIST OF FIGURES

FIGURE	PAGE
1. The Design of the Study	95
2. Wilcoxon Signed Rank Test for the Grammaticality Judgement Test	114
3. Wilcoxon Signed Rank Test for the Elicited Imitation Test.....	117
4. Wilcoxon Signed Rank Test for Task Assessment.....	120

CHAPTER 1. INTRODUCTION: PPP AND TBLT

Overview

The present chapter first provides the background information about history of language teaching methodology leading to Present, Practice, Produce (PPP) and then Task-based Language Teaching (TBLT). Afterwards, the chapter examines the theoretical foundations underpinning TBLT, its advantages, and its criticisms. Finally, the notion of Task-based Language Assessment is explained and its relevance to this study is highlighted.

1.1. Background of the Issue

Communicative language teaching (CLT) emerged in the mid-1970s in the UK as a reaction to the previous methods of language teaching which considered language as a pure linguistic system through which the phonological, lexical and grammatical aspects of language were emphasized (Ellis, 2003; Samuda & Bygate, 2008). In essence, CLT attended to the functional/notional model of language and attempted to develop communicative competence (Hymes, 1972) in language learners. In other words, the focus of CLT was on the ‘use’ as opposed to the ‘usage’ of language, which is the ability to use language meaningfully and communicatively (Ellis, 2003). Samuda and Bygate (2008) contend that by ‘use’ it is meant that the target language is not just there for the purpose of practicing or gaining dominance on, but, more importantly, for the purpose of conveying information, be it for personal, professional, social, political, or artistic purposes. In order to promote language use in the context of Second/Foreign Language Teaching, teachers need to employ holistic activities, one of which is tasks (Samuda & Bygate, 2008). Holistic activities require language learners to engage their knowledge of different component and

areas of language proficiency such as a vocabulary, grammar, phonology, and discourse altogether, as opposed to the analytic activities where these components and areas are worked on in a separate fashion. The rationale is that holistic activities are more authentic and closer to real-life activities that students will encounter in everyday life situation.

Communicative language teaching has a weak and strong version depending on the approach that it adopts towards teaching languages. In the weak version, CLT follows the traditional methodology of language teaching, i.e., present, practice, produce (PPP), in order to teach notions and functions of language and seeks to provide practice for language learners via communicative activities such as tasks. As such, the weak version of CLT is similar to task-supported language teaching (TSLT) notion (Ellis, 2003; Samuda & Bygate, 2008) or 'task-based language teaching' (lower case TBLT) notion (Ellis, 2003; Hunan, 2004; Willis & Willis, 2007; Long, 2014). In the strong version of CLT, language is not broken down into structures or notion/functions; it is discovered through the process of communication. In fact, the strong version of CLT is highly critical of the PPP approach, considering the remarkable qualitative gap between the type of language that learners encounter in class and the one occurring in the real-life context (Samuda & Bygate, 2008). The strong version of CLT is similar to task-based language teaching (TBLT) notion (Ellis, 2003) and Task-Based Language Teaching (upper case) (Long, 2014).

Task-based Language Teaching and the concept of task came to existence and popularity in response to CLT's inefficiencies in that, despite its functional syllabus, it was still following a cumulative view of language where its communicative activities revolved around discrete, pre-selected items where students would go through presentation of items in a controlled manner and ultimately practice them in free production. It goes without

saying that this was the same model as PPP but with the difference that the syllabus was focused on functions rather than linguistic forms. While tasks yielded CLT the opportunity to obviate this problem and get closer to its theoretical principles (Samuda & Bygate, 2008).

Task-based Language Teaching takes language learners' needs analysis as its starting point in order to determine the target tasks that language learners would ultimately need to master using the target language. After the target tasks are identified, they make up the content of the task syllabus. In TBLT, a task syllabus incorporates a number of pedagogical tasks that are supposed to create the same context for language learners as that of the real-life situation. Therefore, tasks play pivotal roles in the design, implementation, and evaluation of a TBLT educational framework (Long, 2015).

1.1.1. TSLT versus TBLT

Long (2015) defines Task-Based Language Teaching (TBLT) as “an approach to course design, implementation, and evaluation intended to meet the communicative needs of diverse groups of learners” (p. 5). In fact, TBLT sets teaching languages in a real and authentic environment as its goal. Task-based Language Teaching has been developed and introduced by researchers of second language acquisition and language educators in response to the teacher dominated and form-focused methods of language teaching in classrooms (Van den Branden, 2006). Task-based Language Teaching holds that the prominent element in designing language curriculum, lessons plans, and even assessment, must be a ‘task’ (Samuda & Bygate, 2008; Ellis, 2009). More specifically, tasks have been defined as “the real-world activities people think of when planning, conducting, or recalling their day” such as responding to e-mail messages, making a sales call, attending

a lecture or a business meeting, to name but a few (Long, 2014). According to Ellis (2003, 2009), in order for an activity to be called a ‘task,’ four criteria must be met. First, the primary focus of students’ attention must be on the meaning to be conveyed, that is, the message they want to communicate. Second, there must be a gap in the task. A gap is the actual problem that the learner is to solve while performing the task. In fact, a gap in a task is what leads learners to the outcome of the task, such as a need to convey meaning. Third, language learners must rely on their own resources, both linguistic and non-linguistic, in order to be able to complete the task. That is, they are free to choose and apply whatever resources that they have in mind in order to perform the task. Last but not least, there must be a clear outcome, preferably a non-linguistic one that language learners achieve when they complete the task. Unless learners achieve the outcome, the task would not be considered successfully accomplished. The current study adopts the Ellis definition of ‘task.’

TBLT holds that tasks should constitute the main component of language learning curriculum. In fact, TBLT considers tasks as both necessary and sufficient in the development of language learning curriculum. However, TSLT, Task-Supported Language Teaching (Ellis, 2003; Samuda & Bygate, 2008), considers a supportive role for tasks in the more traditional approaches of language teaching. That is, tasks are regarded as a necessary but not a sufficient component of language curriculum in TSLT. In other words, tasks do not have a rudimentary role in the curriculum but are just there to provide a means through which the language features that have already been taught could be used in a communicative fashion to consolidate learners’ learning.

It should be noted that TBLT is not a method of language teaching since it does not put forth a series of detailed techniques of teaching. Rather, TBLT is considered an approach of language teaching (Ellis, 2009, 2012). An approach refers to the theories of language and language learning which underpin the practices of language teaching while a method is the overall plan for the presentation of language material in accordance with a certain approach (Richards & Rodgers, 2014). Nevertheless, there is no single version of TBLT but a plethora of approaches to TBLT, each of which stresses an important characteristic of communicative language teaching and learning as applicable in their respective, local contexts. That is to say, TBLT is realized in many different ways around the world. Ellis (2009, 2012) distinguishes between Long (1985), Skehan (1998), and Ellis's (2003) versions of TBLT in terms of the following criteria: a) natural language use, that is, how similar is the context of language use to that of the real-life situation? b) learner-centeredness, i.e., the extent to which students are in charge of their activities c) focus on form, e.g., direction of language learners' attention to form during classroom communication d) the type of tasks, whether it is focused or unfocused. From among these criteria, two are agreed upon among the three versions of TBLT: the provision of opportunities for natural language use and the focus on form. Even so, there is not a consensus among these TBLT versions as to how focus on form should happen. Long (1985, 1990) believes that it should happen through corrective feedback, mainly in pre- and post-task phase, and Skehan (1998) emphasizes pre-task planning and design of the task in provision of focus on form opportunities, while Ellis (2003) contends that focus on form can happen during all phases of task implementation, even in the during-task phase,

through the use of different techniques of focus on form. Furthermore, Ellis (2003) does not consider learner-centeredness as an essential criteria for TBLT.

It is important to discuss ‘focused’ versus ‘unfocused’ tasks when distinguishing TBLT and TSLT. Ellis (2016) refers to focused tasks as activities that are designed to elicit a predetermined linguistic feature in which students are required to use that particular form in order to convey their meaning and complete the task. On the other hand, unfocused tasks do not necessitate a particular target form, thereby giving the language learners the option to choose from among a plethora of language forms in order to get their meaning across and therefore accomplish the task. In terms of the type of task in TBLT, Skehan (1998) only accepts the unfocused tasks in TBLT while Long (1985, 1990) and Ellis (2003) believe in the inclusion of both focused and unfocused tasks in TBLT. To elaborate more, the emphasis on the form and meaning of the target language differentiates between the focused and unfocused tasks. Last but not least, Long (1985, 1990) and Skehan (1998) consider the traditional approaches to language teaching as theoretically unjustifiable and call out for the replacement of these approaches with TBLT, while on the contrary, Ellis (2003) believes that TBLT can work effectively together with the traditional approaches in a modular language syllabus (Ellis, 2012). As with task-based methodology, task types are realized differently depending on the local context.

1.2. Theoretical Rationale for TBLT

Long (2015) puts forth a couple of justifications for the application of TBLT in second language pedagogy. He argues that the premises of TBLT are motivated, supported, and corroborated by the theory and research findings of the past 40 years in the field of Second Language Acquisition (SLA). As an example, TBLT relies on both implicit and

explicit processes as the underpinnings for task implementation which, according to SLA literature, can be used at any age range. The fact that the tenets of TBLT are in line with research findings grants credibility to it. Additionally, the premise of TBLT is deeply rooted in the philosophical principles of education such as student-centeredness, learning by doing, and an egalitarian relationship between teacher and student. As another strong point of TBLT, Long (2015) posits that TBLT, as opposed to the majority of existing second language teaching approaches, does not stand on the extreme sides of using either a synthetic or analytic syllabus. Long argues that this is the main problem of current approaches. A synthetic syllabus puts the language as the locus of attention and breaks the language into linguistic subcomponents such as vocabulary, collocations, or notions and functions, and sequencing them according to the criteria of difficulty or frequency for the language learners. As such, the learner's developmental readiness is not of importance in the synthetic syllabus, while an analytic syllabus takes the learner and the learner's internal process into account. In an analytic syllabus, the language is no longer the object of instruction, as in the synthetic syllabus, but the language is the medium of instruction. An analytic syllabus presents the language learners with authentic and natural samples of L2 where the language learners should analyze this input and induce the grammatical rules. Therefore, the analytic syllabus pays more attention to the message and pedagogy rather than the language (Long, 2015).

The approaches that adopt an analytic syllabus use only a focus-on-forms method and the approaches that use a synthetic syllabus use focus-on-meaning method. The problem associated with the approaches using the synthetic syllabus is their lack of compatibility with the natural processes of language learning (Long, 2015). On the other

hand, the problem with analytic approaches is their lack of enough attention to the importance of intentional learning and their inefficient addressing of learners repetitive grammatical errors. Nonetheless, TBLT offers a remedy by adopting an analytic task syllabus while making use of focus on form methods that could obviate learners' persistent grammatical mistakes. In essence, TBLT offers the opportunity to incorporate the intentional learner which proves effective in dealing with learners' grammatical shortcomings. Provision of negative corrective feedback or a brief explanation of grammatical point in the pre-task phase can be a good example of TBLT's acknowledgement of the power of intentional learning and focus on form (Long, 2014). Long considers focus on form a methodological tenet of TBLT.

Long further argues that learner centeredness has a significant role in TBLT in that the course content is decided through a thorough examination of learners' needs so that the course content could be in conjunction with learner's communicative needs. Furthermore, the form-focused feedback which is provided to students is reactive. That is, the focus on grammatical aspects of language comes in response to learners' errors, which in turn, would be in line with learners' internal syllabus and thus based on their genuine needs. Along the same line, individualization of instruction in TBLT sets to respect individuals' differences in learning, which can be taken care of through relevance of course content with the communicative needs of students. Last but not least, functional language ability is an important goal of language learning which is paid a special attention to in TBLT. Functional language ability allows L2 learners to be able to undertake real communicative tasks such as ordering food for delivery on their phone or giving directions to a passerby.

In the next section, the theoretical underpinnings of TBLT, interaction, input, output, focus on form, and needs analysis which further provide the conceptual rationale for the use of TBLT are discussed.

1.2.1. Interaction, Input and Output Hypothesis

The interaction hypothesis has been first put forth by Long (1980, 1983, 1989), who holds that interaction is conducive to language acquisition. In fact, interaction sets the grounds for the learners to have negotiation of meaning, which in turn, fosters the comprehensible input that is essential for the process of learning. During negotiation of meaning in classroom, students try to make meaning by attending to and adjusting the linguistic features. As such, the language they receive becomes more comprehensible. In fact, the resulting comprehensible input is an outcome of the negotiation of meaning during the interaction which occurs between the learners (Long, 2014; Ellis, 2003). Comprehensible input and negotiation of meaning are in line with the premise of input hypothesis which states that students learn a language through input which is comprehensible and includes the structures at the next stage of students' language acquisition.

Additionally, another outcome of interaction is output modification. In essence, as a result of interaction between learners and the concomitant negotiation of meaning, learners are pushed to modify their utterance in order to make it more comprehensible. Here is exactly where output hypothesis gains significance. Swain (1995) states that when learners are given opportunities to produce their utterance during the interaction, they get to realize the linguistic gap between what they want to say and what they actually are able to say. Therefore, not until learners are given the chance to produce the language will they

be able to notice the gap in their linguistic repertoire. In the development of methodologies for the teaching of languages, TBLT is argued to pave the way for the actualization of interaction through which negotiation of meaning makes the input more comprehensible for students and also allows them to notice the linguistic gap in their output.

1.2.2. Focus on Form

One of the strong points of TBLT is that it is capable of engaging both focus on meaning and focus on form at the same time (Van den Branden, 2006). Long and Norris (2000) state that TBLT attempts to compensate for the shortcomings of the use of focus on meaning in the analytic syllabus through the use of focus on form. In essence, focus on form enhances the rate of language development, which is slow in pure meaning-focused approaches. Additionally, focus on form remarkably improves the accuracy of learners' grammatical proficiency. Many scholars (e.g., Ellis, 2003; Long & Norris, 2000; Skehan, 1998) believe that TBLT has great potential to enhance learners' language acquisition through empowering teachers to design tasks which would increase the probability that language learners' attention would be directed to particular aspects of the language code in the context of a meaningful activity. It is precisely through this—with the task as the vehicle with which to do so—that second language acquisition is theorized to be maximized.

Long (2015) defines focus on form as a reactive approach through which a broad range of pedagogical procedures are utilized in order to direct learners' attention to the linguistic forms in the context of communication problems arising during task performance. He states that focus on form raises the chances that learners' attention to code features will be synchronized with the learner's internal syllabus, developmental stage, and

processing ability. Long further argues that pure focus on meaning, through which learners' learn implicitly by simply being *exposed* to language communication, would not be sufficient for learning a vast majority of language forms especially those non-salient ones and would take an excessively long time to master those forms through mere exposure. Hence, in order to seek the middle ground between the two extremes of the traditional methods of language teaching, which dwells excessively on grammar through focus on formS, and the time-consuming focus-on-meaning approaches, which fall short of improving learners' grammatical proficiency, Long holds that the focus on form, as a reactive method, provides language instructors with the vital apparatus to fine-tune learners' linguistic proficiency through briefly drawing their attention to linguistic code features. It is important to highlight that this all takes place through the vehicle of tasks, task-based interaction, and learners' using the language creatively, employing their own linguistic resources.

Ellis (2003) contends that there are two ways to incorporate focus on form in TBLT: a) through a proactive approach, i.e., the use of focused tasks and; b) a pre-active approach, i.e., through a focus of form methodology. In the proactive approach, the teacher has already set a plan to direct students' attention to a certain linguistic form. Yet, in the pre-active approach, focus on form is the methodology incorporated into the performance of an unfocused task, i.e., the task which does not dwell on any specific forms. In other words, the teacher can direct students' attention incidentally to the linguistic forms which cause problems in their communication. It is recommended in the TBLT literature that teachers employ an eclectic mix of both, depending on the needs of their students, the task, and the context (e.g., Baralt & Morcillo Gómez, 2017).

1.2.3. Needs Analysis

Long (2015) states that in order for a language course to be rational, relevant, and successful, it must take into account learners' real-life goals and their present and future needs. He further points out that a genuine TBLT program necessitates the allocation of resources for a needs analysis and preparation of materials that are in line with the needs of the target population. In fact, a needs analysis is one of the main reasons for which TBLT can be regarded as learner-centered. Therefore, the first step in TBLT course design is to undertake a task-based learner needs analysis, which Long argues should be conducted by competent applied linguists and off-duty teachers rather than by TBLT teachers (Long, 2015). What a task-based learner needs analysis does is target the communicative language needs of learners according to which the program design and delivery will be fine-tuned. (See Long, 2015, for a detailed guide on how to conduct task-based needs analysis and to employ diverse methods and sources). Tasks in TBLT should be chosen on the basis of their congruency with learners' needs; in other words, tasks are considered units that derive from needs analysis in TBLT (Long, 2014). Van den Branden (2006) also stresses that tasks could also be referred to as kinds of activities that learners want or have to be able to do with the new language they are learning. Task-based needs analysis identifies target tasks that are in line with learners' current or future needs. Target tasks which include the real-world activities that people do in their everyday life provide a good example of target discourse samples. An example of target tasks for language learners who are to launch their study-abroad program could be enrolling in classes, reading an academic journal and asking for directions. From this, the researcher or course designer derives target task types, and then, established pedagogical tasks that must be sequenced in a syllabus. Sequencing

should be informed by cognitive complexity (see Baralt, Gilabert, and Robinson, 2014, for a detailed explanation of different cognitive complexity models for task-based sequencing).

Van den Branden (2006) categorizes learner needs as *objective* and *subjective*. Objective needs can be obtained by parties other than the learner by means of the analysis of learner's personal characteristics, and their language use choice, as well as their level of proficiency. An example of objective needs can be the case of a would-be pilot trying to secure a job through mastering the sort of language which would assist him in that regard. Subjective needs are determined through learners' personal statements which might not necessarily be the same as the objective needs. In fact, subjective needs relate to what and how learners *want* to learn. Ideally, in a task-based course, there should be a balance between the objective and subjective needs in using tasks in TBLT (Van den Branden, 2006).

1.3. Criticisms of TBLT

Task-based Language Teaching has received extensive attention since the late 1980s by both second language researchers (e.g., Ellis, 2003; Long, 2014; Skehan, 1998a, 2011) and teacher educators (e.g., Prabhu, 1987, Nunan, 2004; Willis, 1996). Ellis (2009) states that TBLT has gone beyond the level of a theory and evolved into actual practice by stressing the use of language based on function as well as the experiential learning. The special point about TBLT is that it has both theoretical and practical grounds. Task-based Language Teaching started out as a set of pedagogic principles in the applied linguistics literature (e.g., Long, 1985). Since then, it has evolved to be a full-fledged foundation upon which entire language programs are designed. There are now multiple publications on

TBLT, an entire book series dedicated to TBLT (John Benjamins Publishing), a rigorous research conference on TBLT held biennially, and new courses, programs, and even government-funded language training that are implementing task-based ideas. Several countries now mandate that foreign languages be taught in a task-based way. In 2015, the International Association for Task-Based Language Teaching (IATBLT) was founded (www.tbtl.org). Nonetheless, not surprisingly, TBLT, like any other language teaching and learning approach, does not go without criticism (e.g., Sheen, 1994, 2004; Swan, 2005; Seedhouse, 1999, 2005; Widdowson, 2003), owing to the fact that it has targeted the general dominant views about second language acquisition.

In essence, TBLT questions the structural syllabus of language teaching and the premise that language can be broken into small grammatical parts rather it emphasizes, among other things, the significance of the context of language use and interaction resulting from it. Ellis (2014) believes that the root cause of the criticism targeting TBLT is the false assumption that deems TBLT as a ‘single method’ rather than a general ‘approach’ to teaching languages. Ellis (2009) points out two major reasons for the criticisms towards TBLT: the theoretical rationale for TBLT has been misrepresented, and the differences among the proponents of TBLT has been well considered.

Along the same line, Ellis (2009, 2014) puts forth 12 main misconceptions about TBLT, which will be thoroughly discussed in this section. One of the misconceptions regarding TBLT is that *task* is an ill-defined construct where its definition is not clear enough for one to differentiate it from other types of language teaching activities (Ellis, 2009, 2014). For instance, Widdowson (2003), while taking account of Skehan’s (1998) definition of task, holds that the criteria in what Skehan calls the defining features of task

in 'loosely formulated' (p. 26). Widdowson believes this would in turn make the delineating lines between task and other types of activities vague and unclear. It should be noted here that Skehan (1998) considers four criteria for a 'task' including a) meaning is primary, b) there is a goal that needs to be worked towards, c) the activity is outcome-evaluated, and d) there is a real-world relationship (p. 268). Widdowson's (2003) criticism is partly true in that he claims Skehan is not clear whether he means pragmatic or semantic meaning by using the term meaning is his definition. Additionally, his criticism concerning the ambiguity of the term *goal* and the nature of *real-world relationship* seems to be justifiable. Nevertheless, Ellis (2009, 2014) dismisses Widdowson's criticism about the outcome of the task where he states that the successful outcome could be achieved through little language use, which might not lead to language learning. Ellis argues that the purpose of the definition of a task is not to mention the nature of task outcome but to determine what type of educational activities tasks are. Ellis further argues that the misconception of Widdowson results from his generalization taken from only a single definition of task, i.e., that of Skehan's, while Widdowson should consider a wider range of definitions such as Ellis' (2003, 2009), Samuda and Bygates (2008), and Long's (2014) definitions.

As for the criticism on *task*, Seedhouse (2005) argues that *task* cannot be considered as a unit for language courses since tasks might take different performance features according to the context that they are being used, which, in turn makes predictions about the activities resulting from *task* performance next to impossible. Ellis (2014) responds to this criticism, while acknowledging the fact that this statement might be partly true, arguing that Seedhouse is magnifying the issue. Ellis holds that in the first place the predictability of the activity resulting from task only matters in task-supported language

courses where *task* is used to teach the structural syllabus, while, this might not be the case with the TBLT courses where creating the context of language use and incidental learning matters. Ellis further points out that *task* can be designed and implemented in such a way so as it would be possible to predict the nature of the activity resulting from it.

Another criticism of TBLT is what some perceive to be insufficient grammar coverage. Sheen (2003) holds that TBLT lacks a grammar syllabus as TBLT's treatment of grammar is based upon the unexpected grammatical problems that hinder communication, and Swan (2005) takes it as far as stating that TBLT bans grammar. Ellis (2009) contends that the emphasis on grammar depends on the number of focus and unfocused tasks used in TBLT which hinges very much on the type of task-based syllabus used. In other words, TBLT can incorporate a *pure* task-based syllabus, a grammar-oriented task syllabus, or a hybrid task syllabus. The pure task-based syllabus includes only unfocused tasks where the focal attention is on the language use. The grammar-oriented task syllabus makes use of focused tasks where grammatical points are taught through the framework of task. And the hybrid task syllabus which can consist of both focused and unfocused tasks. Ellis believes that Sheen and Swan most probably have addressed the pure task-based syllabus when criticizing TBLT. Ellis (2009) further argues that teaching should not just be seen through the perspective of syllabus since the methodology, which is the way a syllabus is actualized, plays a more significant role. He contends that methodologically grammar has the potential to receive attention at different phases of task implementation, i.e., the pre-task, during-task and post-task phases, even though there is not a consensus among TBLT scholars as to which phase grammar should be worked at. Hence, regardless of syllabus and methodology, TBLT deems an important position for

grammar even though it might not receive the focal attention. In Long (2015), it is emphasized that explicit grammar correction can and does have a place in TBLT, as long as it is reactive in nature (focus on form). These criticisms highlight once again the lack of uniformity in how TBLT is realized across the globe.

Swan (2005) contends that TBLT is appropriate for the second language contexts or what he refers to as ‘acquisition-rich’ environments, whereas, in the foreign language contexts, which he calls ‘acquisition-poor’ environments, TBLT will not work effectively on the grounds that more of a structural approach is needed to enrich learners’ grammatical repertoire for communication. The criticism is intertwined with the hypothesis that unless learners at the beginners’ level are provided with grammar, they would not be able to communicate. Ellis (2009) responds to this critic by putting forth two arguments. First, he rejects the contention that TBLT necessitates production right from day one. He further argues that TBLT can be input-providing and at the same time output-prompting. In fact, TBLT can be input-providing with beginners by emphasizing the Listening and Speaking skills. Ellis states that the input-providing TBLT has the capability to not only account for learner’s ability to comprehend but also endure the grammatical knowledge learners will need in production skills, that is, Speaking and Writing. Second, Ellis believes that beginners do not need grammar to be able to communicate. He points out that the utterances at the beginners’ level are nominal, context-dependent, and scaffolded; that is, learners need help in order to make these utterances. The process of mastering grammar is a gradual and dynamic process which TBLT accounts for. Besides, conversely, TBLT can better serve *acquisition-poor* environments where there is not much chance of communication. In fact, TBLT compensates this lack of communication context through

which learners can have ample interactions, while, at the same time, learners' grammatical resources are built gradually in the context of language use. Ellis (2012) states that TBLT's approach towards grammar might be different from the mainstream approach, however, this does not mean that grammar receives no attention.

Task-based Language Teaching has also been criticized for its learner-centeredness and group-work nature. Swan (2005) claimed that TBLT's emphasis on group activities has pushed teachers, who should be an important source of information for learners, to the margin. Swan contends that TBLT teachers' role as the manager and facilitator of communication is not an efficient role as teachers are the important providers of target language knowledge. Along the same line, Carless (2004) criticizes group work in TBLT as he observed the learners use L1 instead of L2 while performing tasks. Ellis (2012) argues that the assumption that TBLT essentially necessitates group work is not true as there are other ways of task performance, individually and in pairs. Swan should take into account that TBLT teachers' role is not passive since as TBLT teacher can be the manager and facilitator of communication but at the same time be actively engage in the process of providing input and feedback to learners. In essence, TBLT teachers might feel the need to provide brief explicit explanations about form after task implementation. Therefore, a TBLT teacher's role in classrooms is not limited to a specific one but a multitude of roles. Considering Carless (2004), it should be noted that the fact that learners used their L1 instead of L2 is the result of TBLT teachers' derelict of duty not TBLT's inherent problem. As mentioned above, TBLT teachers should actively engage in monitoring the group work of learners in the class, especially in the during-task phase, in

order to make sure that learners are clear about how to perform the task in group and are on the right track.

1.4. Task-based Language Assessment (TBLA)

In the early 1990s, in line with the development of TBLT, a line of research emerged that emphasized task-based assessment. John Norris has been one of the leading scholars in the field on task-based assessment and evaluation of learning, at the lesson, unit, and program levels. The main principle of these alternative assessments was to figure out some other way to assess learners' abilities which would be different from that of the large-scale testing (Norris, 2016). The movement in the testing community — from portfolios to performance testing — set about emphasizing the examination of the ability of test takers to use knowledge, more specifically language, rather than the assessment of rote memorization by discrete items. In line with the advent of performance assessment in mainstream education, Task-based Language Assessment (TBLA) (Mislevy, Steinberg, and Almond, 2002; Norris, 2002), also known as Task-based Language Performance Assessment (TBLPA) (Bachman, 2002), emerged. Task-based Language Assessment has been defined by Brindley (1994) as the process of evaluating the quality of learners' communicative performances as obtained through their goal-directed, meaning-focused language use. Task-based Language Assessment focuses on how well learners can mobilize their language in order to achieve their real communicative goals, as opposed to examining their knowledge of language systems and structures (Brindley, 2013).

Assessment *tasks* play a pivotal role in design of tests in TBLA as they necessitate that test-takers use the language through combining both skill and knowledge and get involved in meaningful language communication (Brindley, 1994; Norris, 2002).

That is to say: while TBLT espouses teaching with tasks, it also espouses assessing with tasks. Ellis (2003) defines assessment tasks as tools that are used in the context of meaning-focused, goal-directed language use in order to elicit and evaluate communicative performance of learners. Assessment with tasks, in fact, is quite different from discrete-skills assessment (DSA) which measures learners' linguistic competence through the use of discrete and decontextualized test items. While TBLT holds that the successful use of language in the social contexts is not bound to the linguistic competence, TBLA, along the same line, takes into account other competencies such as sociolinguistic, strategic, as well as discourse competencies in its measurement of learners' successful language performance. It also, critically, looks at whether or not the task was performed successfully. In other words, TBLA transcends the scope of language knowledge and assesses learners' ability to apply language knowledge appropriately and effectively (Mislevy, Steinberg, and Almond, 2002). The reader is encouraged to look at Nielson (2015, 2016) for an example of an entire task-based program with a full task-based assessment for the implementation of a Chinese online learning program, which was funded by the U.S. government.

What distinguishes TBLA from other types of assessment is that the *measure* of a test takers' performance is inherent in the task itself and is not separate from the task. In fact, what makes a test not be a task-based test is exactly this undue separation. As an illustration, take the example of a test requiring subjects to listen to a lecture and subsequently answer multiple-choice items. The test is indeed performance-based in that it requires learners to process a real-life lecture. Now, is this test task-based? The answer is *no*, given that the measure of language performance is not included in the task, i.e., the

lecture that testees listen to, but it is indirectly and separately measured by the assessor via the analysis of the performance itself, which in this case is the test takers' answer to the multiple-choice items. From the example, it is evident that the measure of students' proficiency must be inherent and incorporated into the performance of the task; otherwise, any form of a separate performance measure would question the task-based nature of a test. In essence, the directness of measurement is a distinguishing feature of TBLA rather than the performance-reference feature (Ellis, 2003).

There are two main procedures to TBLA in terms of assessing task performance: a) the assessment developed from the underlying ability or construct, or b) the one derived from the holistic performance of the target task (Bachman, 2007; Brindley, 2013; Ellis, 2003; Long, 2014). The underlying ability or the construct procedure utilizes test tasks to provide information about learners' mastery of a certain underlying ability or construct of language. In fact, the tests used in the ability-or-construct branch of TBLA are system-based and draw upon a psycholinguistic view of language use in designing tasks. The underlying ability or the construct approach draws upon linguistic components such as grammar, vocabulary, or fluency in order to assess the task performance of learners. Nonetheless, Long (2015) contends that defining the underlying ability or construct is a vexing issue of this procedure. The other TBLA approach, which Ellis (2003) refers to as a work-sample approach, is more concerned with how learners can perform the task and what they can do rather than with their language ability. As Long and Norris (2000) put it, the whole task is the construct to be assessed rather than just the learner's linguistic proficiency. Brindley (2013) holds that it is the 'real-world' criteria which should be taken into account rather than the language criteria in assessing students' performance (p. 1). For

instance, the ability to give a lecture necessitates more elements than the linguistic elements, which could be taken into account through assessing non-linguistic elements as well. More specifically, one can be grammatically absolutely accurate but not able to convey the key point in giving a lecture. Thus, task accomplishment plays a significant role in this particular TBLA approach. Ellis (2003) states that the test tasks based on this approach are often performance-referenced tests. That said, it is important to acknowledge that Long (2015) argues that this approach suffers from the transferability issue; that is, it is difficult to predict real-life task performance based on a certain task-assessment performance.

1.5. The Difference between TBLT and TBLA

Ellis (2003) states that there are two major differences between TBLT and TBLA in terms of the choice of task and the measurement of task performance. As far as the choice of task is concerned, the tasks both in teaching and testing should well cover and represent the construct or the domain being focused; however, the importance and weight given to this criteria differs between TBLT versus TBLA in that the choice of task is a more sensitive job in TBLA than TBLT, owing to the fact that it has direct effects on validity, which determines the credibility of scores. In essence, it is imperative that the task used in TBLA represent and cover the construct and the domain to be tested, while in TBLT it is desirable and satisfactory if the tasks to be taught have this feature. Additionally, the measurement of task performance in TBLA should be more explicit than TBLT, where developing such measures in task-assessment performance to ensure reliability and validity is a painstaking process requiring time and expertise. Hence, it can be concluded that TBLA is a longitudinal investment.

1.6. Advantages of Task-based Language Assessment

Brindley (2013) puts forth a couple of the advantages for TBLA, with particular attention given to classroom-based assessment. He states that TBLA directs teachers and students' attention to using language as a tool for communication rather than having them focus on the language knowledge as an end, which is the case with most of the traditional methods of language testing. Additionally, "TBLA integrates learning process and assessment through the use of *attainment targets* which are directly linked to course content and objectives" (Brindley, 2013, p. 2). TBLA also sets the grounds for learners to receive diagnostic feedback as they can compare their task performance with the clear performance criteria which is presented to them. Furthermore, TBLA utilizes various forms of reporting the assessment outcome in terms of performance which is comprehensible to non-specialists. This would foster the communication between the people who want to use performance information and the educational institutions (Brindley, 2013). Norris (2016) argues that TBLA provides the opportunity to examine multiple aspects of language ability and development such as accuracy, fluency, complexity, procedural knowledge, and pragmatic proficiency through a single performance. He further adds that TBLA can have positive washback effects in that it triggers the educators and teachers to reconsider how teaching and learning happens.

1.7. The PPP Approach

As mentioned above, depending on the functions of tasks in language teaching, two approaches of TSLT and TBLT have been designed. TSLT views language as a set of products that can be mastered in a sequential and cumulative manner. In fact, the main focus of PPP is to elicit accurate target language production from day one (Shintani, 2013).

In the PPP approach, the target language items are first *presented* to language learners through examples that might be accompanied by explicit language instruction. Then, language learners *practice* the target language in a strict, controlled manner through the use of exercises and drills. The goal of practice phase is to help language learners produce the presented material rapidly and easily (Skehan, 1996). Finally, language learners are asked to *produce* the target item in a more spontaneous fashion (though, often this phase can be an extension of the practice). In fact, learners should be provided with the opportunity to use the language freely and flexibly so as they could master the presented material and be able to apply the learned material in a new context (Skehan, 1996). Long (2015) holds that the PPP lesson structure includes the presentation of dialogues and reading comprehension passages which are geared towards the intended grammar of the lesson, then drills and written exercises are intensively practiced, and ultimately students are given the chance to practice more freely through what Long calls “pseudo-communicative language use” (p. 20).

The PPP approach’s relatively long use in the field of second language education and teacher training programs can be attributed to some advantages that it is perceived to offer at the classroom level (Skehan, 1996). The PPP approach puts forth a clear-cut framework where the role of the teacher and the activities to be followed are straightforward and predetermined to the extent that every detail of the interaction in the classroom could be pre-planned and put in the syllabus. In fact, this approach grants teachers a dominant and controlling role in which they have to follow a fixed structure for every session. While this might seem more like a double-edged sword; however, its product is the straightforward syllable that results from the PPP approach leaving no room for

ambiguity or confusion especially on the part of the teacher. Thus, the PPP approach would be more comforting for teachers, especially the novice ones (Skehan, 1996). PPP teacher training programs could also be argued to be less challenging compared to TBLT, owing to the fact that the context of language learning is fixed. In fact, the most challenging part of language teaching is the context in which it is being taught; the more this context is free, unpredictable, and *open-ended*, the harder and more challenging it is for teachers to stay on the track of their syllabus. Probably one of the reasons that many language training programs still stick to the PPP approach at the cost of not facilitating true communicative competency among their studies is the fact that communication, interaction, and the context thereof pose challenges on the straightforwardness and objectiveness of the classroom syllabus, as compared to PPP. As a corollary, the PPP approach can better ensure accountability because testing in this approach tends to be straightforward and the evaluation of the coverage of syllabus is not problematic (Skehan, 1996). Additionally, the PPP approach's underlying theory had strong links with the prevalent theories of language learning. One of these prevalent theories stemmed from the audiolingual method that language learning is a process of habit formation.

Over the course of time and with the advancement of the field of second language acquisition, the PPP approach and its supporting arguments have been criticized. First, the outcome of the PPP approach was to produce the language learners who had poor skills in communication and only certain group of gifted students reached high levels of proficiency through this approach (Skehan, 1989). In fact, the PPP approach failed to meet the high levels of achievement in all four skills. Second, the underlying theory of the PPP approach has been attacked and repudiated. The premise that focus-on-forms and habit

formation can help learners master a language has been devalued through a shift towards more cognitively oriented approaches (Williams & Burden, 2015). Furthermore, the idea that language learners can master a language that has been broken into bits even in the same order that they have been presented with has been questioned in the field of second language acquisition. Skehan (1996) holds that simply presenting the language learners with the language does not guarantee the acquisition as learners' process of internalizing the language is more complex than that. In addition, the PPP approach seems to ignore the role of interlanguage and making mistakes in the learning process. In fact, too much emphasis on the role of accuracy at the cost of losing fluency is another major drawback of the PPP approach.

1.8. Explicit/Declarative versus Implicit/Procedural knowledge

The nature of linguistic competence can be characterized by two different modes of knowledge, i.e., *implicit* versus *explicit* (Ellis, 2005). There is a general agreement among the SLA scholars, that is, both the innatists as well as connectionists argue that second language acquisition involves implicit knowledge. However, how this implicit knowledge is attained and the possible role of explicit knowledge in this process is a bone of contention: What is the nature of explicit and implicit knowledge? And how are they distinguished?

Ellis (2005) puts forth six features that help distinguish these two types of knowledge. He believes that awareness can be deemed as a criterion for this distinction in that explicit knowledge entails the conscious awareness of the linguistic rules while on the opposite implicit knowledge involves the unconscious or, in other words, the intuitive awareness of the linguistics rules. Along the same line, learners can be asked to report after

performing a task or taking a test whether they used *feel* or *rule* in order to complete them-the former corresponding to implicit knowledge, and the latter corresponding to explicit knowledge.

Explicit knowledge is a type of knowledge which is encyclopedic and factual in nature, such as the knowledge of abstract grammatical rules. In order to make explicit knowledge, learners have spent time consciously analyzing the relations and rules between different elements of a system. In contrast, implicit knowledge is automated; that is, learners gain greater control over the rules and fragments of language over time to the degree that little or no conscious attention is needed in order to apply those rules. Ellis (2005) holds that implicit knowledge manifests more systematicity and consistency than explicit knowledge. In essence, he posits that explicit knowledge shows more variability compared to its counterpart knowledge due to the fact that learners engage in tasks with some degree of hunch and conjecture as to how linguistic rules function; therefore, it could be stated that explicit knowledge in learners is more imprecise, inaccurate, and uncertain. On the other hand, even though there is some degree of inconsistency in implicit knowledge as well, Ellis argues that this type of knowledge is more certain in nature when applied in comparison with the explicit knowledge.

Additionally, the processing involved in explicit knowledge is controlled, while in implicit knowledge is automatic as discussed above. Krashen (1982) argues that learners tend to convey their meaning during communication using implicit knowledge which is automatic and quick to apply, while in the meantime if they focus on form, have the related explicit knowledge, and have the time to access them, they would use it to monitor the accuracy of their message. Besides, Ellis (2005) reports that, depending on the tasks that

students perform, they either use their explicit or implicit knowledge. For instance, learners tend to access their explicit declarative knowledge in planning difficult tasks or when performing a think-aloud tasks while performing a GJT. On the other hand, implicit knowledge is applied when fluent performance is required on the part of learners.

Another characteristic of explicit knowledge is that it is verbalizable; that is, explicit knowledge equips learners with explanation of the linguistics rules behind their choice in answering a task. On the other hand, the nature of implicit knowledge is non-verbalizable (Ellis, 2005). The explanation of linguistic rules is fulfilled through the use of metalanguage which is the technical/semitechnical linguistic terminologies to explain linguistics rules. Finally, the last feature in distinguishing explicit knowledge from implicit knowledge is learnability. Learnability means that explicit knowledge can be learned over potentially longer period of time than implicit knowledge. In fact, explicit knowledge can be learned at any age while implicit knowledge is learned at a certain time period due to some limitations that learners face after critical period. Ellis reports that language learners are more conducive to mastering implicit knowledge before their puberty.

Knowing the nature of these two types of knowledge, one might wonder how they can be measured in a test. To this end, implicit and explicit knowledge should be operationalizeable. In so doing, some criteria should be defined as to what characteristics the tests that measure each of these two types of knowledge have. Ellis (2004, 2005) believes that operationalization of implicit and explicit knowledge should draw upon the distinguishing attributes of each of them. As discussed above, the degree of awareness of linguistic rules that a test requires reveals to a certain extent what knowledge that test is assessing. The more the tests are inclined towards learners being aware of these linguistic

rules to successfully take the tests, the more those tests assess the explicit knowledge. It should be noted that this is a matter of degree and represents a continuum rather than an either-or situation. That is, a test can be more predisposed towards assessing one knowledge more than the other.

Time is another yardstick that can be accounted for in distinguishing between the implicit and explicit knowledge on tests. In fact, time pressure in performing a task on-line leaves little opportunity for learners to access their analyzed knowledge of linguistic rules. In timed tasks such as timed GJT, learners resort to using their automatic unconscious knowledge so as to be able to keep up with the time. Thus, Ellis (2005) contends that tests such as timed GJT and oral imitation tests can better assess this knowledge of the learner. Focus of attention is also a criterion that tests aiming to assess implicit or explicit knowledge should take into account. In effect, depending on the purpose of the tests, they can prioritize accuracy or fluency. Accuracy requires focusing on form, while fluency requires production of concepts in order to convey meaning (Ellis, 2005).

Additionally, consistency in response, or what Ellis (2005) refers to as *systematicity*, can indicate whether learners make use of their explicit or implicit knowledge. Therefore, learners tend to be more consistent on tasks that aim their implicit knowledge. On the other hand, tasks that require learners to use metalanguage such as think-aloud tasks tend to measure learners' explicit knowledge. Therefore, systematicity and metalinguistic knowledge are two features that can be utilized in order to receive implicit and explicit knowledge respectively.

Having discussed the distinguishing features of explicit and implicit knowledge, I would like to refer to the relationship between implicit and explicit knowledge. Basically,

there are three main positions as to the nature of the relationship between implicit and explicit knowledge. These are the non-interface position, the weak interface position, and the strong interface position (Ellis, 2005). The non-interface position holds that implicit and explicit knowledge are formed through different acquisitional mechanisms. The non-interface position also argues that the two types of knowledge are held in different areas of brain and are also retrieved for task performance which draw on different processes including the control processing for the explicit and automatic processing for the implicit knowledge. The extreme version of the non-interface position repudiates the idea of explicit knowledge converting to implicit knowledge or vice versa. The weak interface position generally holds that it is possible to convert explicit knowledge into implicit knowledge yet under certain conditions of *when* and *how*. There are three main stances in the weak interface position depending on those conditions. The first stance holds that explicit knowledge can be converted to implicit knowledge conditioning that learners are developmentally ready to master the linguistic form through practice (Ellis, 2005). The second stance contends that explicit knowledge can be effective in an indirect way in the acquisition of implicit knowledge through making specific linguistic forms more noticeable for learners. The third position holds that explicit knowledge can provide the output that functions as the auto-output for the implicit knowledge processes. The strong interface position holds that explicit knowledge can be converted to implicit knowledge through practice, which means learners first learn the declarative knowledge then through rehearsal and practice achieve the level of automaticity and turn it to implicit knowledge; however, there is no consensus among scholars concerning the nature of the practice that turns explicit knowledge, whether it is communicative or mechanical (Ellis, 2005).

1.9. Definition of Key Terms

Task: A task is an activity with a non-linguistic outcome that requires language learner to rely on their linguistic resources in the context of meaningful language use in order to complete the task. In order for students to complete the tasks and reach the non-linguistic outcome, the task should include a gap, which is what students try to fill in order to get to the outcome of the task (Ellis, 2003).

Task-based Language Teaching: ‘(a.k.a. TBLT, task-based instruction, task-based learning): Task-based Language Teaching is a teaching approach that stresses the role of communicative and instructional task as its pivotal unit of instruction. Task-based language teaching advocates the use of tasks for creating a communicative and interactive learning setting through which negotiation of meaning happens among students. In this approach, the linguistic forms of language are acquired through the interactions that necessitate authentic language use (Richards & Schmidt, 2010).

The PPP approach: The language teaching approach which stands for Presentation, Practice, and Production, each of which refer to one phase of instruction delivery, especially grammar instruction. In the Presentation stage, new information is presented and explained. In the Practice stage, also known as the repetition stage, the new information and items are rehearsed either individually or in groups. During this stage, attempts are made to help language learners practice the items in a less controlled fashion. Finally, in the Production stage, also known as transfer stage, learners get to use the practice items freely in a more automatic manner with little help from the teacher (Richards & Schmidt, 2010).

Focus on form: Focus on form is a procedure which including a variety of reactive pedagogic strategies to direct language learners attention to linguistic features in the context that they are used during the communication. Focus on form increases the chances of synchronizing language learner's attention to linguistic features with their internal syllabus, stages of development, and processing ability (Long, 2014).

Discrete-point tests: Discrete-point tests are a type of tests that measure knowledge of single language items, each focusing on one aspect of linguistic form. As an example, a grammar test with sections on adjectives, verb tenses and propositions is considered a discrete-point test. Basically, Discrete-point tests are grounded in a theory that language can be broken into different parts such as grammar, pronunciation, and vocabulary and therefore could be assessed separately. A concrete example of discrete-point tests is a multiple-choice test (Richards & Schmidt, 2010).

Task-Based Language Assessment (TBLA): Task-based Language Assessment is an evaluation process which sets out to assess the language learner's quality of communicative performance. Task-based Language Assessment sets the groundwork for language learners to be assessed through integration of skills and knowledge that requires language learners to have a goal-directed, meaning-focused language use (Brindley, 2013).

CHAPTER 2. TASK-BASED LANGUAGE ASSESSMENT

Overview

Chapter 2 reviews the paradigms of language assessment and what have laid the foundations of Task-based Language Assessment. At the end of this chapter, the purpose and significance of the study as well as the research questions will be presented.

2.1. Paradigms of Language Assessment

An examination of the different paradigms of language assessment would be fruitful in gaining better insight about TBLA. There are three main paradigms in language assessment: the psychometric tradition in testing, integrative language testing, and communicative language testing (Ellis, 2003). The structural linguistics and psychometrics testing methodology informed psychometric tradition introduced in the early twentieth century. Psychometric language tests stressed the significance of objectivity and consistency in measurement; therefore, closed type tests such as multiple-choice tests were given prominence. Not only was objectivity and consistency important in the format of the test, but the analysis of test score would go through various statistical procedures in order to obtain reliability and validity. The role of structural linguistics in this paradigm was to determine the content of the tests, which consisted of discrete elements of language that were tested in light of the four language skills: Listening, Speaking, Reading, and Writing. Psychometric tradition of language testing has its own drawbacks in that it puts too much emphasis on reliability, objectivity and generalizability of results at the cost of neglecting construct validity. That is to say, there is not a strong link between the performance in a test and a certain language proficiency theory (Ellis, 2003).

As for the integrative language tests, the main premise they are based on is that language proficiency or faculty is unitary rather than discrete and multidimensional as believed by the advocates of psychometric testing. In essence, psychometric and integrative language tests share commonalities in stressing the importance of reliability and objectivity but differ in the way they consider the language faculty. In integrative language testing, this language faculty was named as “pragmatic expectancy grammar” (Ellis, 2003, p. 281). The integrative language testing holds that the scores from the grammar and vocabulary tests were highly correlated, indicative of the fact that these tests were actually measuring the same construct, that is, the language processing ability which draws on both linguistic and non-linguistic context. Therefore, this language faculty could be better assessed through holistic and unitary tests approaching real-life language activities. An example of these tests could be cloze-tests and dictation. Even though these claims have later been criticized in terms of conceptual and empirical aspects, this did not stop their use and the research trend examining the design and use of cloze tests, since according to the integrative language testing, the generalization of the results of integrative tests to the real-life performance is tenable (Ellis, 2003).

As opposed to psychometric and integrative language tests, communicative language testing, at its early phases, ignored reliability, objectivity and validity by placing the emphasis on the significance of human subjects in the tests. In fact, at the early stages of its development, communicative language testing regarded the notions of reliability and objectivity as subordinate to face validity. Additionally, the learner’s overall task accomplishment would receive a score rather than linguistic knowledge or language skills. In its later developments, communicative language testing took account of reliability and

construct validity of the tests. To date, communicative tests have three main features. First, communicative tests emphasize performance, which means that test performance and criterion performance should be closely matched. This, in fact, necessitates the use of tasks. Second, communicative tests are authentic; that is, the task assessment input should not be simplified and should be as close as possible to the real-life tasks. Third, the measurement or scoring in the communicative tests is done depending on the achievement of real-life outcome. In other words, the criterion of passing the test is whether the learners accomplish the tasks by obtaining the outcome of the task. Hence, it can be concluded that, among these three main paradigms of testing, communicative testing better incorporates TBLA (Ellis, 2003).

2.2. Performance Assessment

Nearly all language tests have some degree or element of performance in them; therefore, it can be noted that some versions of performance assessment have long been used by teachers in language testing in certain formats and types. As a matter of fact, the degree to which a test is performance-based would better be shown along a continuum of the least direct and authentic one to the most real-world or direct one (Norris, Brown, Hudson, & Yoshioka, 1998). For instance, teachers have used essays to assess the writing ability of students, or have used oral interviews to assess learners' speaking and listening skills, which can be regarded as somewhat direct performance assessment. However, owing to the lack of a terminology for the performance assessment tests, they were referred to as integrative tests for many years. The reason for choosing this label by the testing community was that integrative tests did not have a discrete-point format; that is, these tests would not break the language into discrete parts such as vocabulary and grammar in

their assessment. Contrary to that, they would integrate grammar, vocabulary, pronunciation, and even cohesion and coherence. Additionally, these tests, such as interviews, were, in practice, integrating language skills: Listening, Speaking, Reading, and Writing. Nonetheless, performance tests such as compositions and oral interviews were actually quite different from integrative tests such as cloze tests and dictation, in the degree of emphasis on authentic language use (Norris, et al., 1998).

Performance tests involve some distinguishing criteria that should be accounted for. These criteria draw the delineating line between performance tests and other tests. Norris, et al. (1998) state that the first criterion in performance tests is that the test taker must perform tasks. The second criterion is that the task itself must be as authentic as possible, that is, as close as possible to the real-world tasks. And finally, the success or failure in achieving the task outcome must be rated by qualified raters or assessors. It should be noted that performance assessment, which is typically based on tasks, can either use closed tasks, where they can either have a predetermined objective outcome, or open tasks where there is more than one certain less objective outcome to the task (Norris, et al., 1998). Brown (2004) contends that not all performance tests are task-based tests while all task-based tests are performance tests. In essence, he states that the definition of performance assessment is broader and involves task-based tests. He argues that “any discussion of performance assessment will necessarily include some discussion of TBLA, but the reverse will not necessarily be true” (Brown, 2004, p. 92). Brown holds that there are some instances of performance tests such as compositions and oral interviews which are not TBLA. In his viewpoint, what distinguishes the performance assessments that are task-based from the ones that are not, is that success in performing the tasks has a pivotal

role in TBLA. That is, the construct which is stressed in TBLA is the performance of the task itself rather than the linguistic assessment of the performance which is obtained through the use of task. Indeed, the success or failure of task performance does not matter in the non-task-based performance assessment, so long as the test yields the assessor a performance to score in terms of the linguistic characteristics in order to come up with theoretical and/or pedagogical decisions, which is not the case with the TBLA. Therefore, it can be argued that performance in the non-TBLA is important to the degree that it could provide some linguistic clues about test takers abilities. In other words, performance in the non-task-based performance assessment is a means to an end while in TBLA, the performance has the central role, and it is considered an end itself (Brown, 2004).

2.3. Task-Based Performance-referenced Tests

As target tasks get the focal attention in the TBLT programs so as to enable students to accomplish the target tasks they will undertake in the real life, so is the case with TBLA. In essence, the target tasks will be used as a part of achievement tests that serve to gauge students' proficiency resulting from a TBLT program (Long, 2009). Achievement tests in TBLA programs assume that some forms such as task-based performance tests are criterion-referenced. Task-based performance tests in TBLA, rather than focusing on language as a goal and an accomplishment indicator, concentrate on the successful completion of the target tasks. In other words, the indicators of success or failure in these types of assessment is that the students exhibit the type of behavior necessitated by the needs analysis, which has been previously carried out in order to identify the target tasks and the resulting pedagogic tasks in TBLT. To illustrate, take the example of the target task of ordering pizza on the phone. As long as the students are intelligible enough

to order the pizza on the phone, the task is accomplished; therefore, having students to take a grammatical judgement test of some discrete sentences such as “I’d like to order a pepperoni” would not be regarded as task-based performance test. Robinson (1996) argues that task-based tests should follow performance-referenced assessment; that is, achievement should be measured based upon how the learners perform real-world tasks. In essence, the knowledge of language should be indicated through its use.

2.3.1. Assessment of Task-based Performance-referenced Tests

Now this question might arise: To what extent should we compromise between the task completion and language ability of the student? The answer depends on the purposes of the assessment and the ultimate uses that it would be put to. In this regard, there are different approaches specifying the procedures for the analysis and/or evaluation of task performance (Norris, et al., 1998). One approach championed by some scholars (e.g., Long, 2009; Long & Norris 2000; Mislevy, Steinberg & Almond, 2002; Robinson & Ross, 1996) holds that insofar as the students’ performance does not hinder the communication and the student is capable of accomplishing the task, the student would be able to pass the performance test. Here is an example: If the target task is to make a reservation at a restaurant, so long as the test takers are able to actually do the reservation and communicate their requirement to someone, they manage to accomplish the task regardless of the number of grammatical problems they might have. Hence, the successful accomplishment of the task can be assessed and evaluated by observing the outcome. This outcome-referenced approach is very common in occupational performance testing in vocational training programs in order to issue certifications which has inspired language programs task-based exit testing.

Outcome-referenced testing approaches bear their own problems in the field of second language pedagogy. Norris, et al. (1998) contend that outcome-referenced task-based testing does not yield useful information and feedback regarding certain aspects of tasks that show to be somewhat more difficult. Additionally, they believe that this type of testing approach does not show the “efficiency” with which the learner uses the language for the completion of task (p. 54). Norris, et al. further states that outcome-referenced testing does not leave any room for different outcomes to reach success, which might be the case in the performance of real-life tasks. Generalizability of the results is another important problem of outcome-referenced testing which can be obviated through system-referenced testing (Norris, et al. 1998; Robinson, 1996).

On the other hand, there are some TBLT programs which might emphasize the accuracy of production and penalize students’ performance based upon their grammatical, sociolinguistic, and pragmatic mistakes. But this raises a big issue of the nature and importance of errors. That is, in the case that a test taker manages to complete the task but with grammatical, sociolinguistic, or pragmatic mistakes, how can we specify objectively how many and what type of mistakes are acceptable? More importantly, if a certain objective criterion is set for the number and type of mistakes in the accomplishment of the task, then this might run the risk of turning a TBLT program to focus more on language as an object, since, regardless of the completion of the task, language might become the object and goal, while this need might better be met through focused tasks and focus on form, rather than focus-on-formS, and the corrective feedback which can be incorporated into TBLT courses (Long, 2009). Long (2009) contends that if there is going to be a language accuracy consideration in TBLA, it had better be at the overall macro-level rather than the

micro-level of accuracy. That is, the test takers' general language skills should be holistically rated, having incorporated their task completion.

Additionally, as mentioned above, task-based performance tests will also be criterion-referenced; that is, the purposed of these tests is not to compare students' performance with each other but to assess student's ability relative to a certain criterion. Simply put, if the target task is to successfully tell the difference between two pictures, as long as a student is able to identify the differences and meet the threshold level of the criterion, the task is accomplished and the performance of the student is not assessed in comparison with other students. Therefore, TBLA incorporates task-based, criterion-referenced performance tests as its medium of assessment.

2.4. Task-based System-referenced Tests

The issue of generalizability of performance-referenced tests or what Long (2009) calls *the issue of transferability of learners' ability to the real world*, and the problem of uncertainty on how to group and classify the tasks in order to make sure that the pedagogical tasks and real-life tasks are of the same type, had scholars reconsider system-referenced task-based testing which stresses the importance of the "psychological construct" with the task rather than the completion of the task (p. 55). System-referenced tests deem language mastery as a psychological construct irrespective of its use. These tests are devised to evaluate whether learners, for instance, understand certain words, or have the scanning and skimming ability within a certain time limit (Robinson, 1996). Robinson (1996) states that the advantage of the system-referenced tests is their generalizability to different test sample as well as their easiness to construct and administer, while, their disadvantage is that they lack face validity and do not seem to be authentic. The main

reason for this disadvantage might be due to the fact that system-referenced tests evaluate components of language in a discrete, fragmentary manner as opposed to holistically consider procedural skills. Nonetheless, system-referenced testing can make up for the lack of generalization issue of performance-based testing by placing the emphasis on testing the command of construct or some component of the learners' language ability that can be transferred to real life situation irrespective of how much the pedagogical task and the real-life task might be different (Long, 2009; Norris, et al. 1998). Hence, so long as pedagogical tasks and real-life tasks share the same underlying construct, the generalization could be more firmly claimed. It should be noted that performance-referenced tests have high levels of face validity due to the fact that they resemble the future real-life situation to a great extent. According to Robinson (1996), it would be beneficial if both of these two approaches could be incorporated into the TBLA programs. Robinson further states that integration of these two approaches into TBLA could provide opportunities to make use of generalizability of system-referenced testing and face validity of performance-referenced testing.

2.5. Direct vs. Indirect Tests

Both system- and performance-referenced tests can have direct or indirect modes, which have to do with the relationship between test performance and criterion performance (Ellis, 2003). Robinson (1996) states that procedures of the direct tests are equal or close to the criterion or target procedure, while indirect tests procedures are 'abstractions' or artificial versions of the target criterion procedure. Ellis (2003) contends that direct tests set out to attain a holistic and contextualized sample of learners' use of language. However, in indirect tests, contextualization receives little importance and, as a

result, indirect tests might seem less authentic and more artificial compared to direct tests. In indirect tests, the criterion performance is analyzed and then broken into the components and linguistic features thereof; these components and specific linguistic features would make up the measure of the test. As an illustration, the number of blanks correctly answered in a cloze test might be taken as the indicative of learners' proficiency and further, representing the criterion performance of real-life tasks.

Intertwining the concepts of direct and indirect tests with performance- and system-referenced testing would yield four different categories of assessment, i.e., direct and indirect assessment as well as system-referenced and criterion-referenced assessment. (Baker, 1989; Ellis, 2003; Robinson, 1996). The direct system-referenced tests include the use of a language sample in order to demonstrate a skill (Robinson, 1996). An example of direct system-referenced tests would be traditional tests such as oral interviews, written compositions, and tests which include the transfer of information. Direct system-referenced tests draw on tasks. Indirect system-referenced tests follow the psychometric and integrative tradition. They aim to assess the knowledge of specific aspects of system through multi-itemed tests. A typical example of indirect system-referenced tests would be tests that use multiple choice format about vocabulary and word formation. The direct performance-referenced tests are task-referenced and holistic just like the direct system-referenced tests; however, these two types of tests differ in that the type of tasks used in direct performance-referenced tests attempt to get as close as possible to the real-life situation; that is, tasks in performance-referenced tests are more authentic in that they aim to either assess the actual communicative performance of learners in the real life situation or utilize a simulation of real-life tasks. Simply put, the main purpose of direct

performance-referenced tests is the exact simulation of criterion performance. An example of this type of assessment would be specific purpose language ability such as the ability of a pilot to make conversations with a control tower. Indirect performance-referenced tests do not aim to match test performance with the criterion performance. In this type of assessment, the criterion performance is subdivided into more subtasks or component steps and these subtasks are assessed separately. Therefore, these tests are analytic in the design and sample performance of certain skills. An example of indirect performance-referenced tests would be tests of academic language ability such as TOEFL and IELTS. The summary of this categorization is shown in Table 1 (Baker, 1989; Ellis, 2003; Robinson, 1996).

Table 1. Modes of Language Assessment (Baker, 1989; Robinson, 1996; Ellis, 2003).

Mode of assessment	System-referenced	Performance-referenced
Direct (Holistic)	<p>Traditional tests of general language ability:</p> <ul style="list-style-type: none"> ✓ Oral interview ✓ Written composition <p>Information-transfer Tests:</p> <ul style="list-style-type: none"> ✓ Information-gap ✓ Opinion-gap ✓ Reasoning gap 	<p>Specific purpose tests:</p> <ul style="list-style-type: none"> ✓ Simulation of real-world tasks ✓ Tests based on actual performance of real-world tasks
Indirect (Analytic)	<p>Discrete-item tests of linguistics knowledge:</p> <ul style="list-style-type: none"> ✓ Multiple choice vocabulary or grammar tests ✓ Error-recognition tests <p>Integrative tests:</p> <ul style="list-style-type: none"> ✓ Cloze test ✓ Dictation 	<p>Tests measuring specific aspects of communicative performance in a discrete manner:</p> <ul style="list-style-type: none"> ✓ Tests of specific academic skill

2.6. Measurement in Task-based Language Tests

Owing to the fact that tasks do not inherently provide a *measure*, it is the performance of learners on the tasks which should be examined in order to assess their language proficiency. In this regard, there are three main approaches towards performance assessment in tasks: a) direct assessment of task outcome, b) discourse analytic measures, and c) external rating (Ellis, 2003).

Direct assessment of task outcome depends on the task used in TBLA. In essence, tasks can be of closed or open type, where direct assessment can be done through the use of closed tasks. Closed tasks, which are more objective and leave no subjective judgment on the part of the assessor, have a fixed outcome; thus, the performance in closed tasks is either right or wrong, as opposed to open tasks such as oral interviews which do not have one certain outcome and are, therefore, subject to the assessors' personal judgment. The main criticism towards the direct assessment which is usually observed in the direct-performance referenced test tasks is that it is not clear which it assesses language proficiency as opposed to other abilities of learners in completing the task, even though this issue is not evident in direct system-referenced tests.

The second approach uses discourse analytic measures, which examine features of learners' discourse in their performance of task assessment. These features can be indicative of learners' grammatical, sociolinguistic, discourse, and strategic competences, all of which are part of what is called communicative competence. The learners' grammatical competence can be examined through the measures of accuracy, complexity, and fluency. Sociolinguistic competence is the knowledge of the rules and regulation of language use in an appropriate way in social contexts. Learner' ability to understand and

use communicative function in sociolinguistic context can be an indicative of this knowledge. Discourse competence can be shown in their successful judgement of the rules of cohesion and coherence in all the four language skills. And finally, strategic competence is reflected in students' use of techniques to compensate for their linguistic deficiencies.

External ratings are similar to the direct assessment of the tasks in using the assessor that observes a task performance and makes the judgement. However, they differ from the direct assessment in terms of the nature of the judgement, that is, as opposed to the direct assessment where the judgement is objective, the external ratings involve subjective judgement. External ratings make use of scales that determine the competency or what is being measured and the level of performance in that certain competency. The level of performance in external ratings is usually determined through bands (Ellis, 2003).

2.7. Issues in Task-based Language Assessment

Issues in TBLA may be categorized as either theoretical, such as the issues of validity and reliability, which are somehow common to all sort of tests, or practical, such as the restrictions implementations thereof (Brindley, 2013; Wigglesworth, 2008). One of the most important issues in the field of TBLA is that of authenticity. On the surface, it might seem the mere fact that a task assessment matches a real-life task guarantees the requirement of authenticity in a task. However, it is not the case, as there is more and deeper aspects to this issue. Bachman (1990) holds that in terms of authenticity, both the situational and interactional authenticity should be accounted for in test tasks. The former refers to the extent to which a task assessment matches the real-world task, i.e., the testing context, the latter refers to the extent that task assessment engages the test in the performance. Ellis (2003) contends that taking account of authenticity in test tasks is not

an easy job. As for the situational authenticity, it might be challenging to ensure the equivalence between the test tasks and real-life tasks. Additionally, concerning the interactional authenticity, Ellis (2003) and Brindley (2013) maintain that there is an inherent problem in testing situations that the discourse is going to be artificial as test takers know that they are in a testing situation which will automatically affect their discourse or what Ellis calls a 'test genre construct' (p. 306), which does not simulate the kind of discourse occurring in real-life situations. Additionally, in line with the topic of authenticity, there is the issue of whether TBLA should elicit the best possible performance of learners or it should obtain the performance which is representative of a real-world task is a point of contention in the field of TBLA. That is, should task-based tests be administered with supports such as provision of planning time in order to elicit the best performance of the learners? Or should they be administered in a test setting closer to real-life situation without provision of performance advantage such as planning time? Both approaches have their own advocates either arguing for giving the learners enough advanced preparation or for providing the tests setting and conditions as close as possible to the real-life situation. It should also be noted that this issue is not specific to TBLA and it is present in other types of assessment.

Generalizability is another thorny issue in TBLA. Ellis (2003) defines generalizability as the degree that test performance can be predictive of performance in the real-world situations. In other words, the concept of generalizability targets the validity of tests in that it attempts to measure the level of confidence with which decisions be made about the test-takers based upon their performance on the test (Ellis, 2003). In order to understand the issue of generalizability, two of its aspects should be reviewed: breadth and

specificity, due to the fact that taking account of both of these two aspects of generalizability in a single test is really challenging and difficult to achieve. That is, direct system-referenced tests have the feature of breadth of generalizability in that they measure the learners' language proficiency based upon accuracy, complexity and fluency, which yields the type of result that can be vaguely and generally related to different real-life situations. Hence, the direct system-referenced tests would gain breadth of generalizability, as they can be related to a broad range of situations and needs, at the cost of losing specificity of generalizability. In essence, specificity of generalizability is more taken into account in direct performance-referenced tests where learners' ability in performing a real-life task is assessed. Even though direct performance-referenced tests are successful in achieving the specificity of generalizability, this is gained at the cost of losing breadth of generalizability (Ellis, 2003).

One other challenge facing TBLA is the difficulty in distinguishing the world or background knowledge from the language knowledge of learners in their task performance. Ellis (2003) holds that, even though it seems extremely hard, tasks can still be designed that are, what he calls, "content-fair" (p. 309), that is, to the extent possible content is nearly equally known to all language learners. This issue has been referred to as the issue of inseparability by Ellis (2003). When learners are asked to perform a task where they are required to read a passage about a technical subject, then, it would be difficult to know how to attribute learner's performance success or failure to their specific purpose, background knowledge, or their language knowledge. It should be noted that this problem is not specific to TBLA and is evident in other forms of testing. Practicality is another issue in TBLA as it is costlier and more time-consuming compared to other forms of assessment such as the

traditional discrete-point tests (Brindley, 2013). In fact, what makes TBLA an expensive and demanding method is the means which it requires to ensure the situational authenticity such as training teachers, interviewers and raters, designing test tasks, establishing task banks. Brindley holds that in order to make the best use of TBLA, all of the pre-requisites should be met.

CHAPTER 3. REVIEW OF THE LITERATURE

Overview

This chapter provides a review of the body of research on the differential effects of TBLT and PPP and discusses the gap in the body of research that this study aims to fill. Additionally, the chapter also reviews the role of teacher in previous research.

3.1. TBLT vs. PPP

There have been several studies conducted in order to compare the effects of TBLT instruction with traditional PPP instruction (e.g., De la Fuente, 2006; De ridder et al., 2007; Gonzalez-Lloret & Nielson, 2015; ; Lai, Zhao, & Wang, 2011; Li, Ellis & Zhu, 2016; Shintani, 2011, 2013). De la Fuente's (2006) classroom-based study examined the differential effects of TBLT and PPP on the acquisition of meanings and forms of vocabulary. In addition, the study also focused on the effects of two L2 vocabulary task-based lessons, one of which utilized an explicit, teacher-generated focus-on-forms component (TB-EF), and the other one without it (TB-NEF) on acquisition of meanings and morphological aspects of L2 words. More specifically, the study attempted to find out whether a) TBLT lessons are more effective than PPP lessons in enhancing learning of L2 vocabulary and morphological aspects, and b) whether a focus-on-forms phase at the end of a task-based lesson has a positive effect on learning morphological aspects of L2 words.

Participants of the study were 30 students chosen from a task-supported, communicative Spanish language class. Out of thirty students, nine of them never studied Spanish before, 22 had one year of high school Spanish instruction, and seven had two years of Spanish at high school. The students were presented with the treatment having finished 43 hours of communicative L2 instruction in Spanish. The participants were

randomly assigned to three different conditions. On the first day of the treatment, students were presented with an input-based lesson consisting of two dialogs in order to make them familiar with the context of task so that they would focus more on the targeted words rather than the task context. After the treatment, a pre-test about L2 vocabulary knowledge was administered. On the second day, the treatments for PPP, TB-EF, and TB-NEF lessons were presented. In the presentation phase of the PPP approach, students were presented with the dialog similar to the one they had in their first day treatment with the difference that this dialog used target words. This phase included a focus-on-meaning activity initiated by the teacher, and then directed students' focus on formS. In the practice phase of PPP, the students read the dialogue out loud in order to be able to read the target words and then they did three explicit focus-on-formS activities for 20 minutes. These activities did not create authentic real-life communication. In the production stage of PPP, students were asked to have a role play through which students were given the opportunity to have an output-based, meaning-based activity.

The TB-NEF lesson involved a pre-task, task cycle, and task repetition phase. In the pre-task phase, the same task as that of PPP phase was used. The teacher gave some clarifications regarding the meaning of some words. This phase lasted for 10 minutes and the teacher did not focus on form since the primary focus of this stage was on meaning. During task cycle phase, the students had to perform a role-play information-gap task, where one of the students was a client in a restaurant in Spain, while the other was a waiter. This task had been planned to focus on form and meaning. The planning and reporting phase of task cycle involved having students work two by two to decide about the menu then they had to write it and report to the rest of the class the reason for their selections.

The primary focus in this part was on meaning and students incidentally focused on form. The last phase of TB-NEF lesson was a task repetition phase, where students did free role-plays for 10 minutes. In terms of TB-EF lesson, the first two phase of TB-NEF lesson were used for this lesson two; however, the third phase of TB-EF lesson included two foci of form activities, which were used in the practice phase of PPP lesson. In this phase, teachers explicitly focused on form. Two tests of immediate and delayed vocabulary were administered to students. The immediate test of vocabulary was administered to students right after the task completion during which students were presented with 15 slides of word images and asked to say the words. The delayed test was administered 7 days after the treatment, such that it assessed the retention of target words forms and the acquisition of some formal aspects such as gender and article agreement.

The results of the study indicated that students' retention of vocabulary is affected by the kind of L2 vocabulary lesson they were taught. Specifically, the task-based lessons with a built-in, planned focus on form were more beneficial than PPP lessons since they provided students with more opportunities for negotiation of meaning, output production, and on-line retrieval of target words. De la Fuente concluded by explaining that PPP lessons do not provide that much opportunity for students to produce the target form and PPP seems to be inefficient in directing students' attention on form.

Despite the remarkable findings, De la Fuente's (2006) study had some limitations, such as the limited number of participants and the use of discrete point tests acquisition rather than using a role-play task. Additionally, the use of longitudinal studies lend themselves better to examining the developmental aspect of vocabulary acquisition compared to this type of studies.

De Ridder, Vangehuchten, and Gómez's (2007) study examined the effects of a task-based approach on improving L2 learner's automaticity, which researchers operationalized as a more efficient, more accurate and more stable L2 performance. In essence, their study hypothesized that TBLT is more conducive to higher levels of automaticity in learners' language production compared to the traditional communicative approaches due to the fact that TBLT paves the way for structured repetition and creative transfer of knowledge items. De Ridder et al. also reasoned that TBLT compensates for major drawbacks of the traditional methods by allowing learner-centered activities, and authentic communicative settings, which can have positive effects on the automaticity of learners' language production.

The sample of the study included 68 intermediate-level students of Spanish as a foreign language for Business and Economics at the University of Antwerp. The participants were randomly assigned into two groups of control (35 students) and experimental (33 students). The participants' first language was Dutch and they had all passed the beginners' course Spanish for Business and Economics. Both the control group and the experimental group attended two classes per week over the course of two terms, each twelve weeks long. Both the control and experimental group had to attend four stages of the course. The first three stages were the same where both groups were presented with a strong systematic or focus-on-form components: presentation, explanation, and exercises. The fourth stage was different for the experimental and control group. The experimental group had to attend a total of 10 hours instructions on a task-based instruction called *prácticas comunicativas*. After the four stages, the experimental group shot an advertising spot for a brand new product. The control group did not have instructions. Instead, after

the three stages, they were required to use the previously acquired knowledge in a similar context by reading a passage on Spanish companies extracted from the specialist business press. Afterwards, in order for them to be prepared for the oral exam, the control group was asked to individually gather information for their dossier and to make a short presentation of each of these companies.

The students' performance was assessed, using a six-criterion rubric which was based on the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (De Ridder, et al., 2007). These criteria included pronunciation, fluency, intonation, sociolinguistic competence, lexical competence, and grammatical competence. The study used two raters who video recorded the oral performance of the students to evaluate them on a scale of 1 to 4 for each of the criteria.

The results of the study showed that the control group significantly outperformed the experimental group on pronunciation and intonation, which was contrary to their prediction. However, the experimental group significantly outperformed the control group on grammar (present and past tense morphology and syntax, pronoun use, use of *ser/estar* and *por/para*, prepositions and concordance rules), vocabulary (core vocabulary, 'adjustment to the situation,' phraseology, richness), and social adequacy ('adjustment to the situation,' use of *tú/usted*), which supports the hypothesis of the study. Finally, there was no significant difference between control and experimental group on fluency.

De Ridder, et al. concluded that in terms of fluency, the lack of difference between the control and experimental group can be attributed to the fact that the discourse presented to evaluate the experimental and the control groups was to a remarkable extent prepared in advance, which may be the main justification for the absence of reformulating

phrases, pauses, and fillers that are usually present in spoken discourse and determine the level of fluency. In addition, since the experimental group students had conversations with each other, they did not feel the necessity to improve their pronunciation and intonation as their L2 interlocutors also had accents. However, in the case of the control group, the students had to have an oral exam with the native or near-native evaluator which somehow made them more motivated to adjust their speech to that of the evaluator. Therefore, the conclusion can be drawn that the task-based approach stimulates the process of automatization to a larger extent than a purely communicative course with a strong systematic component. The study's results should be interpreted with some caveats. The study used only two raters who also had the role of course instructors in the study. Moreover, this study did not measure the students' motivation to find out to what degree the task-based students' outperformance might be attributed to the motivating nature of the task.

Lai, Zhao and Wang (2011) examined beginning-level learners and teachers' impression of task-based instruction and the implementation challenges of TBLT in an online course. The study was conducted at a beginning level Chinese classroom at a virtual high school in the United States of America. The online class used asynchronous as well as synchronous platforms. For example, asynchronous activities included e-text self-study, Chinese podcasts, and practice with a Chinese character learning software, all of which were carried out through the course management system, Blackboard. Synchronous activities included meeting with the instructor and other classmates on a weekly basis for one hour through a conferencing system, Adobe Connect. Through Adobe Connect, the instructors and students were able to share and exchange annotated documents. The

researchers designed a TBLT syllabus based on the book that was already being taught at the high school, *Chengo Chinese*.

The participants of the study were 38 beginning level monolingual Anglo-American high school students. Eighteen were male and twenty were females aged 13 to 18. The four instructors in this study (3 males and 1 female) were aged between 22 to 25, none of whom had ever taught on-line or TBLT classes. Interestingly, only two out of the four instructors had previous classroom foreign language teaching experience. According to Lai, Zhao and Wang, in order to decline the risk of having novice instructors in their study, they had all of the instructors have an extensive workshop and debriefing sessions with the researchers prior to the start of the semester. The study used six different sources of data from both teachers and students including:

- A background survey asking students' demographic as well as foreign language learning and online learning experience.
- Weekly reflection blog entries where students wrote self-reflections about how well they had performed, what they had learned, the struggles and challenges they had encountered, and the strategies they wanted to share with their classmates.
- Class observations and recorded synchronous sessions where the researcher randomly observed one class of each TBLT teacher. In fact, all instructors were required to record their teaching sessions each week via the recording option in the video conferencing system. The recording would log every moment of the teaching session including the aural and written teacher-student and student-student interactions.

- Course evaluation which included three Likert scale questions and four open-ended questions asking students about their impression of their class as well as their view about their perceived learning.
- Students' recorded oral performance of a descriptive task during the final exam. In this phase, each student logged into the conferencing system individually and took the final test one-on-one with the instructor, where they had to orally describe a descriptive task for the instructor.
- Weekly debriefing and interviews at the end of the semester. Teachers met with the researcher through the semester to talk about their opinions about TBLT and the challenges they faced in this program. Additionally, an interview was also done with teachers at the end of the semester to obtain teachers' opinions about TBLT.

The researchers did a qualitative analysis of teachers' and students' impression of TBLT as well as a quantitative analysis of the fluency, complexity, and accuracy of students' oral performance in the final task. The results of the study revealed that Chinese students and teachers had a positive view towards TBLT classes. The study also showed that students did not have the required skill and strategy to perform effectively in the TBLT course. It should be noted that the implementation of TBLT was not without its challenges due to the difficulties in designing the TBLT syllabus and issues in the implementation of full task cycle. Additionally, the arrangement of virtual classroom was inflexible and this in and of itself affected the implementation of collaborative tasks. It was also challenging to build rapport among students in the online conferencing system. However, all in all, TBLT proved to be effective in lowering cognitive load of students as well as in fostering students' participation.

The findings of the study revealed that there is a mutually-beneficial relationship between TBLT and learner strategy training. That is, TBLT needs learner strategy training in order to be more effective and at the same time TBLT consolidates the effectiveness of strategy training by fostering autonomous learning among the learners. The learner strategy training includes: a) macro-level training where learners are made familiar with the philosophical, pedagogical, assessment bases of TBLT ahead of the course, and more importantly assisted in harmonizing the TBLT syllabus and the e-textbook for the online course; and b) micro level training where learners are trained to develop metacognitive, cognitive, social, affective strategies which would increase learners' gain from TBLT (Lai, Zhao & Wang, 2011). Metacognitive strategies in on-line TBLT classes refer to the linguistic features that learners need to attend to during text chatting. Cognitive strategies refer to how learners should negotiate meaning and form in an on-line class. Social strategies include training learners in how to build up a rapport amongst themselves during the on-line class interactions. Finally, affective strategies help learners keep motivated despite not having one-to-one contact with the instructor and other students. The other issue in this TBLT on-line program relates to the time limitation for the implementation of the pre-during-post task phases within the short time of synchronous sessions. Lai, Zhao and Wang (2011) put forth the solution of having learners work on input-based tasks before the synchronous sessions and use integrative pre-tasks or review tasks during synchronous sessions to go through the during- and post-task phases.

Shintani's (2011) study examined the differential effects of production-based activities and input-based tasks on learners' acquisition of vocabulary. The participants of the study were 36 Japanese students between the ages of six to eight. The students were

randomly assigned into two experimental groups and one control group. The experimental group classes involved students aged between six to seven having four months experience of learning English, while control group involved students aged between seven to eight year having sixteen months experience of learning English.

The study had three types of treatments: a) an input-based group that was presented with input-based instructions, b) a production-based group that was presented with production-based instructions, and c) a control group that received a set of three activities (English songs, Total Physical Response, and alphabet practice), without being exposed to any of the target words. The instruction time was 30 minutes for all of the conditions and was done by the researcher. In order to gauge vocabulary knowledge, four tests were designed and administered over the course of two weeks. Of the four tests, two were production-based and two were comprehension-based. The test were administered three times as pre-test, posttest 1, and posttest 2. The classroom sessions were audio- and video-recorded.

The target vocabulary included 24 words which were presented to the production-based group at three different intervals with eight words each time in every other lesson. All the 24 words were taught six times for the instruction-based group. The reason for the difference in presenting the vocabulary to the two experimental groups was to prevent students from becoming aware of the goals of research. The tasks used for instruction-based group included three listen-and-do tasks which were repeated throughout all the six lessons. The tasks included: a) Task 1 called, help the zoo and the supermarket, b) Task 2 called help the animals, and c) Task 3: a listening bingo game. The production-

based group was presented with five tasks including listen and repeat, guess the hidden items, throwing dice, production bingo game, and Kim's game.

The study's comprehension tests included a multiple choice listening test during which students were asked to listen to words being pronounced and choose the related picture from among 6 pictures. The test included forty questions, sixteen of which were distractors. The other comprehension test was a category task test during which students were required to listen to sentences and figure out in which context each sentence was stated. There were four contexts, including a fruit and vegetable shop, a kitchen, a bathroom, and a zoo.

The production tests were discrete-item production tests and a 'Same or different' task test. In the discrete-item production test, students were required to name each vocabulary item presented to them through flash cards. Each correct answer would be counted as one point in this test. In the 'Same or different' task test, both the researcher and the student had different sheets with 24 pictures of objects on them. The students were required to name the object and check with the teacher to see if they had different or same objects. If they had the same object, then the student would put a checkmark next to the object; otherwise an x would be used.

The results of the study showed a significant advantage for both input-based instruction and production-based instruction over control group in post-tests 1 and 2 on all tests. A main conclusion from this finding is that the input-based instruction can help young L2 learners acquire L2 vocabulary. Thus, the finding supports the hypothesis that both IB and OP instructions fostered the acquisition of both receptive and productive knowledge. However, the comparison of input-based versus production-based condition did not yield

a significant difference in terms of their effects on acquisition of new vocabulary. In fact, IB and PB groups had similar achievements on three tests of the multiple-choice listening, the discrete-item production, and the ‘Same or Different’ task. In terms of the fourth test, that is, the category task test, the input-based group performed better than the production-based group in both posttests.

Interestingly, the study found out that students got the chance to produce during input-based instruction and comprehend during production-based instruction. That means both conditions fostered comprehension and production vocabulary knowledge. The study concluded that the input-based condition is more effective and provided better opportunities for interaction which might explain the better performance of input-based group in the category task test and equal performance on production in spite of having fewer chances of practicing production. Shintani’s (2011) study has some limitations, such as the low number of participants. In addition, the study worked with children, so this puts into question the generalizability of the findings to other (and adult-based) contexts. The use of students from a private school may also limit the generalizability of the results.

Shintani’s (2013) study set out to investigate the differential effects of input-based focus on form and production-based focus on formS on learner’s vocabulary acquisition. In the present study, focus on formS was operationalized through present-practice-produce (PPP) approach while focus on form was realized via a task-based approach. In focus on formS, the focus is on intentional learning and production, while in focus on form, students are inclined towards the form of language in an indirect and incidental manner and also the meaning is the primary focus. The participants of the study were 45 six-year-old L2 learners of English from Japan with no prior experience of English

language instruction. Learners were randomly assigned into three groups, each including fifteen L2 learners. Out of these three groups, two were experimental groups of focus on form and focus on formS and one was a control group. Each of all three groups were broken into two classes of six to nine people where they met two times per week.

In the Focus on formS condition, five activities were carried out following a PPP approach. In the present phase, the first activity was carried out where the participants repeated some words. In the practice phase, two activities were used where learners were asked to pronounce the words on a flash cards once chorally and the other time individually. In the production phase, the students had to say the name of the object shown on a card in order to win and collect the card. During all of these activities, students had to focus on accurate production.

In the focus on form condition, three tasks were used whose completion could only be possible through understanding the input. The tasks involved the learners listening to the teacher's orders and responding accordingly. For instance, the teacher would command: "Please take the crocodile to the zoo" and the learner would respond by selecting the correct card and putting it in the correct holder.

In order to assess students' performance, Shintani used two tests: a discrete-item word production test and a 'Same or Different' task test. In the discrete-item word production test, the researcher asked students to name the target words written on the flash cards. There were a total of 24 flash cards for this test, on which both adjectives and nouns were tested. As for the 'Same or Different' task test, each student worked individually with the researcher. Students had pictures related to words and adjective numbered from 1 to 24. The objective of the task was to have students check whether their pictures were the

same as or different from the teacher's. For instance, the researcher would ask the student "What color is it?" or "My soap is pink. Is your soap pink?" in order to elicit responses. All of the interactions were recorded.

Shintani found that the focus on formS group significantly outperformed the Focus on form group in both the discrete-item test and the 'Same or Different' task, on both the immediate and delayed post-tests. In fact, students were able to use nouns in both the controlled and free production test. This positive effect of the Focus on formS condition had not been reported for the use of adjectives. The only positive effect reported for adjectives was reported in the controlled production test. In fact, students did not show any communicative mastery over adjectives. The results of the study confirmed the effect of Focus on form conditions to improve students' productive mastery over nouns and adjectives. Therefore, the researcher concluded that students do not necessarily need to produce words to be able to build productive knowledge of the words.

González-Lloret and Nielson's (2015) study took into account the effectiveness of a task-based Spanish program implemented at the United States (U.S.) Border Patrol Academy (BPA). This purpose of the program was to train U.S. agents to better communicate in Spanish in order to help injured people, to communicate with immigrants abused by smugglers, and to calm families of prisoners. The task-based Spanish program was designed by TBLT experts upon the request of BPA due to the fact that previous old courses, which had strict grammatical syllabi, were reportedly ineffective as many of the agents who completed the program were still unable to communicate in Spanish.

In order to measure the efficacy of the new BPA program, three empirical studies were conducted. The main objective of this evaluation was to find out whether agents'

performance on task assessments can be generalizable to other contexts beyond the test context. To this end, one study set out to compare the oral proficiency of students (agents) trained in the previous grammar-based course with the oral proficiency of students trained in the new TBLT course. The second study used a computerized oral proficiency instrument to measure whether or not students' oral proficiency improved over the course of TBLT instruction. The third study was qualitative in nature and explored students' perceptions about the Spanish TBLT program.

The researchers hypothesized that the TBLT group would perform better than grammar-based students on the measures of fluency, lexical complexity and syntactic complexity, while the grammar-based students would outperform TBLT students on grammatical accuracy. The participants of the study included 20 students from the TBLT course and 19 students for the grammar-based group. In order to measure assessment, an oral picture-guided narration task, as well as an audio-recording of students, were used.

The results from the first study showed that the TBLT students performed significantly better than grammar-based students on measures of fluency and structural complexity. However, in terms of lexical complexity, no significant differences were found. In addition, there were not significant group differences in grammatical accuracy. That said, even though the TBLT students' grammar instruction was occasional and contextualized thanks to focus on form, the TBLT students managed to gain grammatical accuracy over time.

The second study, which used a computerized oral proficiency instrument examined the effects of TBLT on students' overall Spanish proficiency. Participants of this study were 256 students who were the first to finish the Spanish Program at BPA. The

study examined whether proficiency level of students is associated with their proficiency improvements. That is, the study hypothesized that the more advanced students' performance would be higher than the beginning-level students; however, both groups would manage to accomplish the task. Along the same line, a pre-test before a posttest after the TBLT course was administered to students. The Versant Spanish test, which is a computer-scored oral proficiency assessment, was used due to its convenience in administering the test and its high correlation with other measures of oral proficiency. The aggregate score on Versant Spanish is a weighted average of the four subcategories (sentence mastery, vocabulary, fluency, and pronunciation). The Versant test has a score scale of 20 to 80. The study hypothesized that the more advanced students' performance would be higher than the beginning-level students; however, both groups would manage to accomplish the task. Along the same line, a pre-test before the TBLT course and the posttest was administered to students. The results of the study showed that the performance of students on the Versant test did change with a mean overall increase of 7.473 points in the posttest—an improvement of 12.5%. In fact, the study showed that the TBLT curriculum was effective not only on students' immediate performance of the task but also on their overall proficiency. The results of the study also indicated that the TBLT course helped both the advanced and beginning-level students improve their overall proficiency. The results also suggested that both the advanced and beginning-level students benefited from the TBLT course and the hypothesis of the study that both advanced and beginning-level students would accomplish the task successfully held true.

The third study, which was qualitative in nature, had the students complete two electronic surveys about the perception and impression of the TBLT program at BPA. The

survey had been sent via email to by the current students enrolled in TBLT course and the students who already graduated and started their job as an agent. Twenty-one students and sixteen agents completed the surveys. The survey included two sections: a four-point Likert scale ranging from ‘Strongly Agree’ to ‘Strongly Disagree,’ as well as two open-ended questions asking students what was it that they enjoyed most about the course and what about the course could be improved. The results of the survey indicated that students were motivated to study Spanish after graduation. They also believed that their TBLT classes concentrated on helping them learn how to do their jobs and the Spanish they learned was very much pertinent to their job needs. The students also believed that, despite their being able to use Spanish to do their jobs, they did not feel confident enough to talk to native speakers outside of the academic context. In response to the two open-ended questions, the students felt that learning of “practical,” “everyday Spanish” to be able to “talk to Spanish speakers outside the job” would be effective (González-Lloret & Nielson, 2015, p. 539). In terms of students who had graduated from BPA, the results indicated that they were satisfied with TBLT program and believed that the program was useful in regards to the topics and vocabulary. The students also added that what they learned through this program was highly applicable to their jobs as new Border Patrol Agents.

In sum, the results of the three studies reviewed here showed that the new TBLT program at the BPA was successful. That is, the students in TBLT program performed better than grammar-based students in oral accuracy, fluency and complexity while they performed equally in grammatical accuracy. Students’ overall proficiency was enhanced, and the course was successful and useful for learners at all proficiency levels. Finally, the qualitative study suggests that students perceive the program as useful in preparing them

to undertake the tasks in their job, although the program was not that successful in helping students communicate beyond their job tasks.

The results of González-Lloret and Nielson's (2015) study must be interpreted with caution as there are a couple of limitations. First, the sample size for the first study was very small. Second, the study could not gather data from the grammar-based course students to compare data for the Versant Spanish test; therefore, the study could not figure out how students would have performed on Versant Spanish test.

Li, et al.'s (2016) study compared the differential effects of task-based and task supported instruction on the acquisition of the English passive structure. The researchers operationalized the three different tasks implementation procedures as: Focus on Meaning, TBLT, and TSLT on students' learning of explicit and implicit knowledge of the passive structure. 150 EFL middle school Chinese students participated in the study. The participants of the study were chosen from five eighth grade classes with 55 to 60 students each. Thirty students were then randomly assigned to five groups: one control group and four experimental groups. The experimental groups were presented with a two hour treatment during which they had to do two dictogloss tasks in which the passive structure was used. The experimental groups had four different instructional conditions: 1) Focus on Meaning (FoM), which performed the two dictogloss tasks without any intervention; 2) The TSLT group, which received pre-task explicit instructions and then performed the tasks; 3) The Focus on form (or pure TBLT) group, which performed the task while receiving corrective feedback on the targeted structure; and 4) the "stronger" version of TSLT, which received pre-task explicit instruction + corrective feedback while performing the tasks. Finally, the control group took a pre-test and post-test only.

The two dictogloss tasks were performed in the following way: the students listened to a story presented to them by the teacher three times, and then they rehearsed it in pairs. Students were then required to retell the story to the class, as well as add their own an ending to the story. Finally, the class was asked to vote for the best story ending. It should be noted that the two narratives included a news report about a car accident and a story from *Reader's Digest* about an earthquake in Haiti. The explicit instruction for the TSLT and “strong” TSLT experimental groups involved a mini lesson about the passive structure which lasted for 15 minutes.

In order to measure learning, the researchers used a grammaticality judgment test and an elicited imitation test. Doing so allowed them to gauge the effects of the treatments on students' explicit/declarative as well as implicit/procedural knowledge. The GJT required the students to specify whether a particular grammatical structure was correct and in case it was not correct, the students were asked to correct the structure. The EIT required the students to listen to the recordings of 35 sentences read by a native speaker, determine if each statement was true, and then had to repeat the sentence in correct English.

Results of the study indicated that there were limited effects for the FoM condition on students' learning of the passive structure. Li, et al. (2016) propose three reasons for why this might have been the case. First, the passive structure is a late-acquired structure and students might not have been at the right developmental and mental readiness to learn it. Second, doing just two dictogloss tasks might not have provided the students with the necessary opportunity to learn the target structure. Finally, in the presence of exclusive FoM, the learners may not have been able to activate the necessary cognitive processes to learn the target structure.

The task + corrective feedback (Focus on form) condition showed some achievement in the acquisition of explicit knowledge; however, this effect faded away over the course of time by the delayed posttest. Li, et al. (2016) concluded that the reason for the short time effect of explicit instruction on students' acquisition of explicit knowledge can be attributed to the short length of instructions presented to students as the instructions were successful in making the explicit knowledge but not effective enough to transfer the explicit knowledge to automated knowledge.

As for the explicit instruction + task (TSLT) condition, the results showed a benefit for the use of explicit instructions before the task on learners' explicit knowledge. As opposed to the TBLT condition, the effects of this condition on learners' explicit knowledge was more durable and did not fade away in the course of time. However, there was no effect on learners' automated knowledge found for TSLT condition. The Explicit Instruction + task + corrective feedback condition (the stronger version of TSLT) showed the strongest effect on learners' explicit knowledge. This condition had the strongest effect on automated knowledge. It seems that the corrective feedback that the learners received improved the effects of the pre-task explicit instruction on student's acquisition of the target structure.

Overall, the two TSLT conditions and the TBLT outperformed the FoM and control group in the acquisition of explicit knowledge. Moreover, the stronger version of TSLT, the explicit instruction + task + corrective feedback, had the best effect on the explicit knowledge; however, it should be noted that none of the conditions outperformed the control group on developing the learners' automated knowledge. In sum, that the

researchers concluded that TSLT is more effective in teaching new grammatical structures than TBLT.

One of the greatest limitations with Li et al.'s (2016) study is the tests that were used to measure learning. A GJT and an EIT cannot be a good measure of students' achievement in the TBLT condition. In fact, TBLT's effectiveness was measured through the use of tests that only assessed students' linguistic competence while linguistic competence is one of the several competences that TBLT affects. The use of these discrete-point tests yields distorted and unrealistic effects of TBLT. In order to measure the effectiveness of TBLT instruction, students' performance on a task performance test must be assessed. As such, researchers can be sure that all the communicative competences of students are being assessed. Most of studies reviewed here indicate a positive effect for TBLT compared to a PPP approach which are shown in Table 2.

Table 2. Summary of Studies Comparing the PPP Approach with TBLT

Study	Target Form	Participants	Treatment	Assessment	Teacher Information	Teacher's Role
❖ De la Fuente (2006)	19 Spanish Target nouns	30 Elementary Spanish L2 Learners	PPP lesson, TB-EF (explicit, teacher-generated focus-on-forms) lesson, and TB-NEF lesson (explicit, teacher-generated with no focus-on-forms)	two dialog tasks	no information provided about the teacher	PPP, TB-EF, and TB-NEF instruction
❖ De ridder, et al. (2007)	pronunciation, fluency, intonation, sociolinguistic competence, lexical competence, and grammatical competence	68 university students of business and economics	a task-based phase called: the <i>prácticas comunicativas</i> was used	oral test and presentation	no information provided about the teacher	giving instructions and feedback

Table 2. Summary of the Studies the PPP Approach with TBLT (Continued)

❖ Shintani (2011)	24 nouns in English	36 Japanese children	Input-Based (IB) & Production-Based (PB) instructions	multiple-choice listening test, Category task test, Discrete-item production test, 'Same or different' task test	the researcher played the role of the instructor	giving instructions and feedback and having interaction with students
❖ Lai, et al. (2011)	fluency, complexity, and accuracy of students' oral performance	38 beginning level monolingual Anglo-American high school students	teaching TBLT syllabus for two weeks	background survey, blog entries, class observations by Likert scale questions, performing a descriptive task, and interview with the teachers	four instructors aged between 22 to 26 who had no prior experience of TBLT were used. The instructors were given intensive workshops on TBLT	teaching TBLT syllabus

Table 2. Summary of the Studies the PPP Approach with TBLT (Continued)

❖ Shintani (2013)	24 nouns & 12 adjectives	45 Japanese beginners of English	Focus on formS (present-practice-produce) & Focus on form (Task performance)	discrete-item word production test & Same or Different task test	the researcher, who had 10 years' experience of teaching played the role of the instructor	giving instructions and feedback
❖ Gonzalez-Lloret and Nielson's (2015)	fluency, complexity (lexical words), and accuracy (noun-modifier agreement and noun-verb agreement)	19 students for grammar-based and 20 students from TBLT course	students were presented with an oral picture-guided narration task using a six-vignette story	performance-based assessment tasks	BPA instructors have been used. No more details about the instructors were provided	giving instructions
❖ Li, Ellis, and Zhu (2016)	English passive construction	150 Chinese middle school English	2-hour treatment session where they performed two dictogloss tasks	Grammaticality Judgment Test (GJT) & an Elicited Imitation Test (EIT)	The instructor was a PhD student with eleven years of experience who had not taught the learners prior to this study.	giving explicit instructions for Dictogloss task & providing explicit feedback

Overall, there are some important points to be taken into account in the interpretation of the results. One of the most important is the issue of whether the tests or tasks used for assessment in these studies can efficaciously measure the particular construct in the treatment. In some of the studies examined, the test or the task chosen to measure learning and the task performance was not a valid tool to assess the particular construct (e.g, Li, et al., 2016). The important point is that TBLT attempts to help language learners achieve communicative competence. Communicative competence incorporates some other competences such as linguistic competence, discourse competence, sociolinguist competence, and strategic competence (Ellis, 1991), which would make it so difficult for traditional discrete-point tests to measure them. Therefore, task assessments in these studies should make sure that they are measuring properly the construct they purport to measure.

As Ellis (2003) stated, TBLT considers language learning as a process as opposed to a product, which is viewed mostly by traditional the PPP approach. Therefore, as a learning process inherently necessitates a course of time, so does its assessment. In fact, assessing TBLT courses and testing need more of longitudinal studies than cross sectional ones. Moreover, the comparison between PPP and TBLT must be focused more on the methodology, which means if there is going to be a more effective comparison, the effects of these two methodologies must be assessed over the same tasks. In fact, the main difference between TBLT and the PPP approach is not about which task to use but how to implement the task and its related methodology.

3.2. The Role of the Teacher

Teachers' roles can vary depending on the methodology they are defined within. In less communicative methods of language teaching, such as the PPP approach, the teacher has a more accentuated role and would even in some cases be the sole speaker. On the contrary, the teacher's role in TBLT is one of modeling collaboration, observing and monitoring students' performance, and intervening when students are facing problems (Ellis, 2003). Ellis even considers the role of a task participant in classroom collaborating with students in their pairs or group to perform the task. The same roles for teacher could be considered when they are participating in a research project. Nonetheless, the issue that has been nearly neglected in many research studies is the use of authentic classroom teacher rather than an external teacher. In fact, as a rule of thumb, to obtain more authentic and pragmatic results, it is imperative that researchers the most extent possible not manipulate the classroom setting for the sake of the research goals. In most cases, the researchers would themselves assume the role of teacher in which case there would be a blow to the authenticity of the classroom, not to mention the issue of bias in the research. In this section, the body of research on the comparison of the PPP approach and TBLT is reviewed in order to examine if they have used the classroom's real teacher or they just used the researcher as the classroom teacher.

De la Fuente (2006) and De ridder et al. (2007), in their comparison of the PPP approach and TBLT, did not report any information on the teacher participants and it is assumed that the teachers in these studies were also the researchers. Shintani (2011, 2013) used the researcher as the instructor used in her study rather than the classroom's original teacher. Gonzalez-Lloret and Nielson (2015) did not provide any information regarding

their instructors nor their training in TBLT. Additionally, Li et al. (2016) used a teacher participant who had not taught the learners prior to the study and, more importantly, provided no information about whether the teacher participant had any prior training in TBLT. Similarly, Lai et al. (2011) used instructors who had no prior training in TBLT. These examples indicate that the important role of the classroom teacher has been taken for granted in light of the research goals. The aforementioned body of literature either used the researcher as the classroom teacher which would undermine the authenticity of the research results (Kim, 2016), or, these researches simply worked with a teacher that had no prior training in TBLT, which could in turn call into question the quality of the instruction that the students received.

3.3. Statement of the Problem

3.3.1. Issues in the PPP Approach

The PPP approach has been criticized on several grounds. First, second language acquisition research has indicated that language acquisition is not just a mere mastering of a set of accurate products in a cumulative fashion, but involves interlanguage development through which language learners gradually move towards accuracy in the process of using language for communication. Second, Ellis (2003) holds that language learning is a process with developmental sequences, some universal. Language learners must go through these developmental stages, following their own “internal syllabus,” as opposed to PPP’s product view of language. Third, it is very difficult, as well as unnatural, to make sure that language learners use the target form in the production phase (De la Fuente, 2006). Finally, the production phase of PPP is very controlled with an exclusive focus on formS and leaves

little leeway for learners to use their own devices, or make their own meaning (De la Fuente, 2006).

Skehan (1996) argues that linguistic achievement rate in the PPP approach is relatively low and most of the students leaving the PPP programs have serious trouble in using the target language. The major outcome of the PPP approach, except for the gifted learners, is relative failure. Moreover, and more importantly, the theory behind the PPP approach has been criticized on several grounds. In essence, the theories advocating focus on form and automatization have lost credibility both in the fields of linguistics and psychology. The most recent theories of language learning contend that the field of language teaching is more than just converting input to output. These theories of language learning accentuate the role of the language learner as the builder of language knowledge, as well as, that of the authentic meaning-based activities where learners have a choice in approaching them. The student-centered and meaning-based class activities are the areas in the literature that has attracted a lot of attention and prompted a good deal of research (e.g., De la Fuente, 2006; De ridder et al., 2007; Gonzalez-Lloret & Nielson, 2015; Li, Ellis & Zhu, 2016; Lai, Zhao, & Wang, 2011; Shintani, 2011, 2013) by delving more into the PPP approach and TBLT via juxtaposing their effects on the quality and quantity of learners' learning resulting from each of these major approaches. That is, the pros and cons of PPP versus TBLT in helping the learners master a language have been the center of attention in a body of research in the field of SLA.

3.3.2. The Importance of TBLA

Long and Norris (2000) contend that in a genuinely TBLA, tasks have a pivotal role of being the fundamental unit of analysis, motivating item selection, test instrument

construction and the rating of task performance. According to Long and Norris, the goal of TBLA is more than just using the real-world task in order to elicit a certain component of the language system which are then measured or evaluated; TBLA, rather, evaluates the performance of the task as the most important construct. In fact, language performance such as fluency, accuracy and complexity can only be considered in the evaluation of task-based performance conditioning that they are inherently related to accomplishment of an assessment task.

The development and implementation of TBLA for TBLT programs follows six steps (Long and Norris, 2000). First and foremost, the intended uses of TBLA within the TBLT program should be precisely specified. As such, four important issues should be taken into account including: who is going to use the information from TBLA, what information the assessment purports to assess, what the objectives of the assessment are, and finally, what the consequences of the assessment would be—that is, what and who is going to be affected by the assessment. Second, target tasks which have been chosen or designed after the needs analysis are examined and grouped according to different task features such as setting, type and amount of L2 used, and number of steps involved in completing the task, so as these features may be replicated in different assessment conditions. Third, test tasks, their formats, as well as their performance evaluation are determined in accordance with the analysis of the task features. Fourth, Long and Norris believe that this is the most important stage where the rating criteria for the task assessment is determined. These criteria include the real-world critical elements which relate to the aspect of task performance, and levels of success in task completion. Fifth, the test task, testing procedure and instrument, and testing criteria must be evaluated in order to ensure

their effectiveness and match with the intended assessment objectives. Finally, the validation of the intended use of TBLA must be evaluated. In other words, this validation process must be ongoing as well as systematic and determine whether the test instruments are providing the reliable and useful information (Long & Norris, 2000).

Task-based Language Assessment consists of tasks working as achievement tests that assess what students have gained from the course. Long (2015) contends that TBLA has two features. First, it involves task performance tests. The focus of these tests is on gauging students' ability to do real tasks rather than on the language itself. Second, these tests are criterion-referenced; that is, the students' abilities are not compared with other students, but the students' performance on that task is evaluated solely based upon whether or not they accomplish the task. In the present study, TBLA will be given a special role in examining students' mastering of the related language knowledge. In essence, the tasks utilized in TBLA will measure communicative competence of language learners rather than just their mere knowledge of the form of the language. In this study, an important point of bifurcation between the PPP and TBLT was the different approaches they adopt in examining students' learning gains and achievement, since TBLA is a relatively newer area in assessment than assessment in PPP, it will be give more emphasis in this study.

3.3.3. Teacher's Role in TBLT

Teachers can assume different roles depending on what teaching methodology they follow. Throughout language teaching history, role of teachers have evolved from being a sole speaker to a facilitator. Ellis (2009) deems a very important role for TBLT teachers despite the fact that TBLT promotes student-centered instruction. In reaction to

Swan's (2005) criticism that TBLT teachers have more of a passive role where they only assume the role of a facilitator of communication rather than a source of information, Ellis (2009) holds that Swan's assumption that TBLT teachers are only bound to be a manager and facilitator of task and communicative activities is wrong. Ellis goes on to state that the role of teacher in providing form-focused feedback during the task performance necessitates a more active role than just a facilitator or manager of tasks. In fact, teachers might at times be prompted to intervene and explicitly teach some problematic language items. Undoubtedly, this requires the teachers to be equipped with both implicit and explicit strategies of providing feedback. Ellis concludes that TBLT has the privilege of having both a student-centered and teacher-centered instruction. Van den Branden (2006) contends that tasks are subject to different interpretations depending on the educational goals, learning needs, and the style of interaction. He argues that despite the fact that in SLA research, tasks have been deemed as a fixed variable where learners are supposedly working on the same task which have rather the same effect on language learning, in fact, tasks are absolutely flexible and can have different effects according to different teachers and students who use it. Therefore, the role of teacher in making the best use of tasks used in classroom becomes more salient. Van den Branden further argues that teachers need to take two important points into account in their practice of tasks in classroom so as the task could provide the opportunity for high quality activities and raise the chances of eliciting actual learning out of these activities. These two actions include a) stimulating learners to put as much mental energy as possible into completing the tasks, b) providing confidence-boosting interaction which would encourage learner's task performance, focus on form, negotiation of meaning, and input comprehension, all of which play a pivotal role in SLA.

According to Van den Branden (2006), teachers should make sure to provide students with not only high quantity input but also high quality input in TBLT. Hence, it goes without saying that great care and exactness must be taken into account in choosing the qualified teachers in task-based instruction and research. Fulfilling these goals through teachers necessitate well-organized task-based teacher-training programs. Nevertheless, despite the importance of teachers' training in TBLT, the previous body of research on task-based research took teachers' roles for granted in that either the background and specialty of teachers delivering the task-based treatment have not been mentioned (e.g., De la Fuentes, 2006; De ridder et al., 2007), the researchers themselves carried out the task-based treatment (e.g, Shintani, 2011, 2013), or a teacher was chosen who had some years of experience in teaching but these researchers did not specify if the teachers had experience specifically in the area of TBLT (e.g., Gonzalez-Lloret & Nielson, 2015; Li, Ellis, & Zhu, 2016).

3.4. Purpose of the Study

Considering the significance of the headings mentioned in the statement-of-the-problem section—issues in PPP, TBLA, and the role of teachers in TBLT—the present study attempted to examine the PPP and TBLT in light of these aspects. In other words, this research study hypothesized that owing to the fact that TBLT sets the grounds for the inclusion of focus on form as opposed to focus on formS promoted by PPP (Long, 1991), TBLT would yield more productive results for the students in that students would have a better mastery of the target feature. In fact, as mentioned above, focus on form allows the students to use the target feature in the context of communication and language use. Long (1991) holds that this would obviate problems associated with focus on formS which tends

to be very boring and full of repetition as well as the problems with the focus on meaning that does not seem to prepare students' grammatical proficiency well enough. Simply put, this research study was an effort to juxtapose the differential effects of focus on form promoted by TBLT versus focus on forms promoted by the PPP approach on students' mastery of explicit and implicit knowledge of the past passive structure in English.

Furthermore, the present study attempted to use performance assessment in its evaluation of students' performance. The reason is that performance assessment and more specifically TBLA allows researchers to assess students' performance using the language rather than just examining their linguistic mastery of a certain target feature. TBLA, which emanates from the tenets of performance assessment, takes students' communicative skills into account. For example, the students of the present study had to plan how to implement the task as well as how to use the language appropriate for that certain context. All of these necessitate an ability beyond linguistic mastery of a target feature. More importantly, there is a dire need to incorporate TBLA when the researchers are evaluating the effectiveness of TBLT in their research (e.g., De la Fuente, 2006; De Ridder et al., 2007; Gonzalez-Lloret & Nielson, 2015; Lai, Zhao, & Wang, 2011; Li, et al., 2016; Shintani, 2011, 2013). Additionally, as opposed to the previous research investigating the possible differential effects of PPP and TBLT (e.g., De la Fuente, 2006; De Ridder et al., 2007; Shintani, 2011, 2013), the present study paid special attention to the choice of TBLT teacher used to provide the treatment. That is, the teacher who assumed the responsibility of providing TBLT treatment is in fact a TBLT expert with several years of teaching TBLT and doing TBLT teacher training, alongside publishing prolific research in the field of TBLT.

3.5. Significance of the Study

Following Sheen (2003) and Swan's (2005) criticism on the lack of enough empirical evidence to support the effectiveness and the possible superiority of TBLT over the traditional focus-on-forms approaches of language teaching such as PPP, there was a need to conduct more studies in order to delve more into the efficiency of TBLT and its possible superiority over PPP. To this end, the present study aimed to build upon the previous literature (e.g., De la Fuente, 2006; De ridder et al., 2007; Gonzalez-Lloret & Nielson, 2015; Lai, Zhao, & Wang, 2011; Li, et al., 2016; Shintani, 2011, 2013) on the possible differential effects of the PPP approach and TBLT on students' language learning in order to examine whether or not these two different approaches in language teaching would yield different effects on learners' mastering of the past passive structure in the English language. Along the same line, not only did the present study examine the TBLT and PPP methods in terms of their instruction effectiveness, but it also used their respective assessment. The use of each methods respective assessment is a point that has been neglected in the previous research on the comparison of the different methods of language teaching (e.g., De la Fuente, 2006; De ridder et al., 2007; Gonzalez-Lloret & Nielson, 2015; Lai, Zhao, & Wang, 2011; Li, et al., 2016; Shintani, 2011, 2013). As an example, Li, et al. used the GJT and EIT, which are typical of the PPP approach assessment to also examine the effectiveness of TBLT. This likely led to distorted results. To obviate this problem in the literature, the present study compared the differential effects of these two language teaching methodologies through using their respective assessment in order to obtain more realistic results.

The assessments used in the study were set to investigate different types of abilities. The GJT and EIT were supposed to measure respectively the explicit and implicit knowledge of students, respectively. In addition, the present study used Task Assessment in order to investigate the students' ability to use the language in the context of authentic language use. In fact, by using these four types of tests, the study could yield a more comprehensive picture of the effectiveness of these methodologies by testing them across the tests, each of which is meant to assess a different aspect of language mastery. The GJT was supposed to assess how well the students were aware of the rule and regulation at work in this grammatical feature; the EIT was supposed to assess how fast the students can recognize and use the target structure, this necessitated that the students have an unconscious and automatic knowledge of the target structure. Last but not least, the assessment tasks were set to examine the students' use of the target linguistic feature in the context of authentic language use. In a nutshell, the assessment used in the present study would hopefully add more reliable and valid results to the body of literature which is nearly dearth of using TBLA in their comparison of the different methods of language teaching. In light of this goal, the following research questions guided the present dissertation:

3.6. Research Questions

The research questions of the study focus on three concepts: a) language teaching methodologies, i.e., PPP and TBLT b) the type of knowledge that language learners master, i.e., the declarative and automated knowledge c) the type of assessment, i.e., the GJT, the EIT, and Task Assessment. Below the research questions of the study are presented in detail.

- 1) Do Present-Practice-Produce (PPP) and Task-Based Language Teaching (TBLT) treatments have different effects on learners' performance on a Grammaticality Judgment Test (explicit/declarative knowledge), as compared to a control group?
- 2) Do the PPP and TBLT treatments have different effects on learners' performance in an Elicited Imitation Test (implicit/automated knowledge), as compared to a control group?
- 3) Do the PPP and TBLT treatments have different effects on learners' performance in a task-based language assessment (true task-based performance), as compared to a control group?

3.7. Limitations of the Study

A few limitations emerged from the study. The present study did not examine the role that corrective feedback could play in improving the effects of the TBLT and PPP methodology on students' language proficiency. Additionally, the effects of the prior explicit instruction on the improvement of students' linguistic performance, especially in the task-based performance, had not been considered. The effects of corrective feedback as well as the prior explicit instruction were not investigated, as opposed to Li et al.'s (2016) study, due to the focus of the present study on the assessment aspects of the TBLT and the PPP approach. In fact, the main focus of the present study was to make up for the previous research's (e.g., De la Fuente, 2006; De ridder et al., 2007; Gonzalez-Lloret & Nielson, 2015; Lai, Zhao, & Wang, 2011; Li, et al., 2016; Shintani, 2011, 2013) lack of inclusion of TBLA in their comparison of the two methods of TBLT and the PPP. Therefore, rather than concentrating on the effects of instructional elements such as the

corrective feedback and prior explicit instruction, the present study took account of the assessment elements.

CHAPTER 4. METHODS

Overview

Chapter 4 provides information about the means of the data collection including the participants of the study, the teachers, the instrumentation, assessments and treatments. The chapter also discusses the pilot studies that helped fine-tune the means of the data collection.

4.1. Participants

The participants of the study were chosen from an English Language Institute entitled Parsian Language School in Mazandaran province located in northern Iran. Parsian Language School is a privately-owned language institute that teaches the English and French languages. The course books covered in this institute for adult English language learners was the book series of Interchange. The study included 18 female and 16 male Iranian English as a Foreign Language (EFL) learners. The English proficiency level of all students had already been assessed by the Parsian Language School upon their entrance into the School as they were obliged to take the placement test prior to the commencement of their English courses. The participants of study had been studying in the Parsian School for at least 3 semesters and were at the intermediate level of proficiency at the time of research. The learners' proficiency was measured by a set of end-of-the semester exams designed by the institute. The participants ranged from 15 to 32 years of age. The participants of the study were sampled because of their availability and convenience of participation. Another reason for choosing the participants of study at this level was to make sure that students had as little knowledge as possible about the target structure of the study, i.e., the past passive voice In English: the students of Parsian School had not been taught the target feature of the study prior to the commencement of this research. The

participants' grouping was based on the class list of the institute, that is, the students were sampled conveniently based upon their institute's class lists. In order to ensure the students' voluntary participation in the study, the researcher required them to read a synopsis, provided by the researcher, of what the study was about and then to sign a written consent form so as to give their consent to take part in the study.

4.2. Teachers

Two teachers of the Parsian institute were used in the present study. Both teachers were in the midst of their doctoral studies in the field of applied linguistics. They both had the experience of taking courses of language teaching methodology as well as Task-based Language Teaching in pursuit of their doctoral studies; therefore, both were familiar with the TBLT and PPP methods. The two teachers were male and in their mid-thirties. They had over 10 years' teaching experience in teaching English as a foreign language in an Iranian context. The teachers were given the necessary explanation and information about how to conduct the study, more specifically the classroom treatments, through Skype connection for over an hour. The teachers were given instructions as to how administer the test and how to do the treatments especially the task implementation according to Willis and Willis (2007) model.

4.3. Instrumentation

4.3.1. Pre-assessment Instruments

Pre-assessment involved a GJT, an EIT, and a task-based test called Task Assessment administered to students at the pre-assessment phase. The GJT and EIT were taken from Li et al.'s (2016) article, as the present study is quasi replication of that study, and the Task Assessment was designed by the researchers (Appendix C). As explained in

the review of their study, the GJT and EIT were designed to assess learners' declarative and automated knowledge of the target structure. The Task Assessment was designed to assess learners' proficiency to use the target structure in an authentic context of language use via the means of a real task. The present study hypothesized that the GJT and EIT would be in line with the assessment premises of the PPP approach and that the Tests Task would be in line with the tenets of TBLA. To put it another way, the methodology of the PPP approach, which follows the psychometric and integrative approaches in assessment, would use the GJT and EIT (Ellis, 2003). In other words, these two tests would be a better reflection of the outcome of the PPP approach. On the other hand, TBLT has its own assessment which follows the performance assessment approach towards testing. The Task Assessment is in line with the tenets of TBLA and performance assessment. In fact, this is where the design of the present study is more robust than Li et al.'s (2016) study by including the Task-based Language Assessment.

4.3.1.1. The Grammaticality Judgement Test

As mentioned above, the GJT has been borrowed from Li et al.'s (2016) study. This test included 40 items whose grammaticality was to be judged by students. Out of those 40 items, 30 items were about the target structure, i.e., the past passive structure in English, and 10 items were distractors; that is, the sentences which used other structures than the past passive voice so that students would not realize the structure which was being assessed. Each item included a sentence with a blank line in front of and a blank line underneath. The instructions of the GJT required students to judge if a sentence was grammatically correct or incorrect. The students were asked to put a 'C' in the blank in front of the sentences that they believed were correct and an 'I' in front of those which

were incorrect. Additionally, the students were also asked to write the correct English version of those sentences they believed were incorrect on the underneath the items. Students received 1 point for answering each item correctly. (It should also be noted that the 10 distractor items were not counted in the scoring of the GJT). Therefore, on the whole, the GJT was scored on a scale of 0 to 30. It should also be noted that the GJT was supposed to assess students' explicit knowledge of the target feature, i.e., the knowledge of the rules and regulations of the past passive voice in English.

4.3.1.2. The Elicited Imitation Test

The items of the EIT have also been adopted from Li et al.'s (2016) article. The test purports to assess learner's automated knowledge of the target structure. In other words, the time taken for students to react to the test item can be indicative of their automated knowledge of the form. The students were required to listen to the recording of a native speaker of English reading 35 items. There was an 8-second time interval between each item, which had been determined through a pilot study after which students' reaction time to answer each item was calculated. The EIT included 35 items, 5 of which were distractors; that is, they were not assessing the target structure and were not counted for the analysis. The EIT required students to listen to each item played by the digital voice recorder and determine whether each of the items was true of their life and then repeat the item in correct English within eight seconds' time interval. As an example, for the item: "My father was hit in a car accident last year," the students had eight seconds' time to first indicate if this actually happened to their father by saying *Yes* or *No* and then repeat that sentence in correct English. The students received 1 point for answering each item correctly; thus, the test was scored on scale of 0 to 30.

4.3.1.3. Task Assessment

The Task Assessment was a narrative task which required the students to read a hypothetical short story about a robbery that occurred in Miami. The narrative used the simple past tense to recount the robbery. The story used the active voice to explain the robbery scene and what actually transpired during the robbery. The instructions of the Task Assessment required students to recount the story using the past passive voice and thereafter add an ending to the story. There was a two-page space underneath the story for the students to rewrite the story using the English past passive voice and add an ending to their own desire. The total number of the sentences to be changed into the passive voice were 17 and students received one point for accurately changing each active voice structure into the past passive voice. Thus, Task Assessment was graded on a scale of 0 to 17.

4.3.1.4. Coding

The coding of the GJT was carried out by considering half a point to each of the two parts of the past passive structure, i.e., half a point was considered for the correct use of the auxiliary verb and another half a point for the correct use of the past participle. The coding of the EIT was done by allotting half a point to each of the two parts of the past passive structure, i.e., half a point was considered for the correct use of the auxiliary verb and another half a point for the correct use of the past participle. Finally, the coding of the Task Assessment was done by allotting half a point to each of the two parts of the past passive structure, i.e., half a point was considered for the correct use of the auxiliary verb and another half a point for the correct use of the past participle.

4.3.3. Treatment Instruments

4.3.3.1. The TBLT Tasks

The first treatment task given to students was a picture task including nine items. Each item of this task included two pictures. The task story was about a hypothetical character named Rebecca and her brother who recently moved away, and a lot has changed since then. Rebecca was going to write her brother a letter explaining what changes she had made since he moved. The task required students to help Rebecca write her letter and explain the changes. Overall, there were nine items each having two pictures with a verb prompt. The students had to look at the two pictures and see the differences and changes that had transpired ever since Rebecca's brother left. Then, using the verb prompt indicated on an arrow, the students had to use the past passive voice in the blank underneath each item to explain what changes had occurred. In order to avoid students' confusion, the subject of the passive structure was printed at the beginning of each blank.

The second treatment task required the students to read a note left by the mother of character of the story, Cindy, asking her to do some chores while she (Cindy's mom) was away for some days. The students were asked to work in pairs and help Cindy write a text message to her mom reporting that she did each one of the items she had been asked to. In order to write this message, the students had to use the past passive voice and the name of each items accomplished in the place of the subject of the past passive sentence. In order to better help students in the process of writing the message, the subjects of the passive voice structure for the 10 items of the task were provided to students so that they could have a clearer picture of how to complete the task.

Since the past passive structure lends itself to explaining the process of making something, a task was chosen that would require students to describe the process of making wheat bread in a traditional fashion. In essence, what makes the past passive structure lend itself to this activity is the fact that the doer of the actions in a process would not be of significance while the changes each item undergoes receives the importance by becoming the subject of the sentence. In this task, the students had to first read a passage about the task characters, John, his father, and their family's timeless tradition of making bread. The story started with John's father explaining to him in response to his question about how bread used to traditionally be made. After reading the story, the students were required to remember the story in order to put the sequence of events in the same order that John's father explained. The students had to work in pairs and number each of the sentence in the blank next to them. Overall, there were 18 items to be numbered understanding of which required comprehending the past passive structure.

4.3.3.2. The PPP Activities

After 20 minutes of instruction about the past passive voice (which will be elaborated upon in the procedure section), the students had to practice the structure through the following activities: a) a discrete item activity including 10 items where the students were required to convert the verb in the parenthesis to the passive voice in order to complete the items, b) a close passage including 7 blanks where students had to read the passage and complete the blanks within the passage by the choosing the correct form between two option provided in the parenthesis, and c) a transformation activity including 10 items of complete sentences in the active voice; the students had to change the sentences into the passive voice. In front of each item in the parentheses, the subject, verb (in its simple form),

and object of the passive voice were presented to students, and the students had to just form the whole sentence by adding an auxiliary verb and the past participle part of the passive structure.

As for the production activities in the PPP treatment, the following activities were utilized:

- 1) A repetition drill where the teacher required students to repeat 10 items, first the whole class and then accompanied with individual spot checks.
- 2) A substitution drill including 10 items where the teacher repeated a sentence in the passive voice and then provided a prompt being a pronoun such as they, him, and so forth where students had to repeat after the teacher by using the prompt pronoun in their repetition and making the necessary changes.
- 3) A restoration activity where the teacher would say three words actually being the subject, verb in the simple form, and the object of the passive voice. The students had to make a full sentence in the passive voice adding an auxiliary and participle to the sentence when repeating it after the teacher.
- 4) A backward build-up activity where the teacher broke a sentence into three parts in their repetition. First, the teacher would use the object of the passive voice and the students would repeat it, then he would add the auxiliary and the past participle in repeating the sentence, the students would do so in the repetition. Finally, the teacher would say the whole sentence and the students would repeat after him. Clearly, the teacher in this methodology of instruction had an active role of making students memorize the grammatical feature.

4.3.3.3. The Control Group Activities

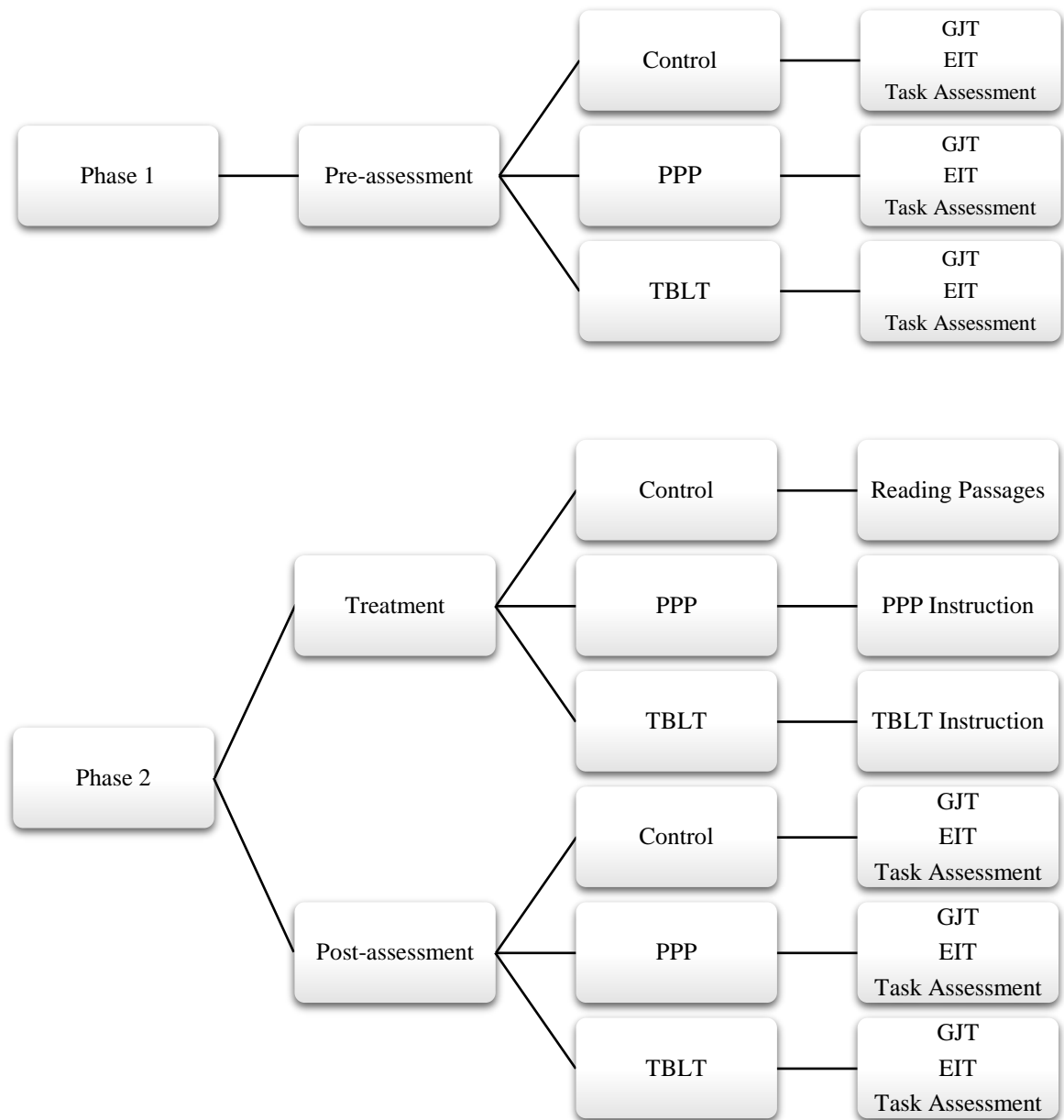
The treatment for the control group consisted of five reading passages taken from a reading book entitled *Steps to Understanding* authored by Hill (1980). The passages were at the intermediate level of proficiency according to the book categorization. Attempts were made to choose the passage that included zero application of the English past passive structure. At the end of each of the five passages, there were comprehension questions that students needed to answer. The stories narrated a funny story which were approximately 15 lines each on average (see Appendix F).

4.4. The Design of the Study

The study used a quasi-experimental design with a format of the pre-, immediate post assessments. The study was conducted in two phases: a) the students were administered the pre-assessment, and b) the students were given the instructional treatment and then were immediately given the immediate post-assessment. The assessments included two tests of the GJT and EIT which measured students' declarative and automated knowledge, respectively, as well as a Task Assessment which aimed to assess students' authentic use of the target language. It should be noted that the GJT and EIT were the same along the pre- and post-assessments except that the order of items were switched in order to avoid practice effect on the tests.

All groups of the Control, PPP and TBLT received the three tests and their performances were compared across the tests. Figure 1 below shows a bird view design of the study.

Figure 1. *The Design of the Study*



4.5. Procedures

4.5.1. Pilot Studies

Before the collection of the data from the pool of participants in this study, the three assessments of the study were administered to two groups of native English speakers and Intermediate-level ESL learners. As for the first pilot study, the native English speakers who agreed to take the tests were 18 Undergraduate students of education at Florida International University attending their undergraduate principles of ESOL course. These students had the goal of becoming the future school teachers in Florida. These groups of students only took the Task Assessment of the study due to the time limitations. While taking the Task Assessment, these students asked about the instruction of the Task Assessment. In fact, the only trouble that they faced while taking the Task Assessment was the use of the technical term *the passive voice* as they did not have a clue what the passive voice meant; however, when the teacher explained to them what the passive voice meant, they did the Task Assessment perfectly well with no mistakes. Afterwards, the test instructions have been modified to avoid the misunderstandings regarding the use of the term *passive voice*.

As for the second pilot study with the English Language Learners, the researcher chose a pool of English Language Learners at English Language Institute (ELI) at Florida International University. Having talked to the director of the institute, Mr. Sanchez, one of the researchers managed to get the permission of the manager and the teachers to do the research at ELI. After the researcher talked to one of the experienced language teachers in the ELI, she kindly agreed to assist the researcher in the data collection, especially in the language lab and introduced the researcher to other teachers. The researcher chose three

classes at the intermediate level to do the data collection. The different classes were randomly assigned to different conditions of TBLT, PPP, and Control. The first class to collect the data from was a TBLT class.

The class was held on Tuesday mornings and the teacher of that class kindly cooperated with the researcher and gave the full time of her class to the data collection. The data collection took around 90 minutes and still some students were struggling to finish before the end of the class. After this class, I coordinated with two other teachers to collect data on Thursday of the same week. However, in the meantime, according to the teachers and management of the ELI, some students of that class complained about the length of test and showed their unhappiness with spending the time of their class on taking tests rather than getting ready for their own institute exam. As a result of this complaint, the researcher could not collect data on Thursday of that week. After speaking with the director of the ELI, the researcher was given the chance to collect the data however, this time under the condition of only using the students' lunch break rather than using their actual class time. In order to do this, the researcher had to go to different classes, talk to students, explain the research, and get the contact information of those who showed willingness to participate in the research. The researcher managed to get the contact information of overall 15 students for one of the classes, and set a time in their lunch break from 12:00 to 1:00 p.m. The researcher also offered to buy their lunch and they unanimously agreed to have pizza for their lunch. Sad to say, on the day of the data collection, only one of the students showed up and the rest were not present to take the test. This was an end to the researcher's endeavors to collect data from the EIT; however, the data collected was sufficient to be used for the pilot study.

The students in the ELI took all the three tests of the study in the language laboratory of the institute. They first took the GJT, then they were recorded doing the EIT, and finally they took the Task Assessment. The students' feedback regarding the test indicated that they believed the tests were long and tiring. Therefore, the Task Assessment was somewhat made shorter by having them write the target structure instead of the whole stories. In addition, the ESL learners seemed to have struggled with the instructions of the tests as most of them did not understand the instructions. Therefore, the instructions of all the three tests were paraphrased with simple words so as to make sure that the students were clear about what each of the tests required them to do. Along the same line, an example was shown at the beginning of each tests in order to help students grasp what the test was all about. In addition, the pilot study confirmed that 8 seconds' interval between each of the EIT items would be enough time to respond and the repeat the EIT items in order to assess their implicit knowledge.

4.5.2. Pre-assessment

The pre-assessment session was held in the Parsian Language. The three classes for the pre-assessment included students who were studying the English language book entitled *Interchange* at the intermediate level in this institute. The three classes were randomly assigned to three groups of Control, PPP, and TBLT by randomly pulling their names out of a hat.

The students of the TBLT group were the first to receive the pre-assessment. The instructor spent a Monday class on administering the three tests to the students of the TBLT group. He explained to the students that they had to judge the grammaticality of 40 items. The instructor ensured the students that there was no penalty for wrong answers. He

then distributed the GJT among the 13 students of the class. The students were asked to start their second test, i.e., the EIT upon their completion of the GJT. The GJT required that students put an 'I' in the blank in front of items which were grammatically incorrect and a 'C' in the blanks in front of the items which were grammatically correct. Additionally, there were blanks under each item of the GJT in which the students wrote the correct form of the items that they thought were incorrect. Out of the 40 items in the GJT, 10 of which were distractors. Those 10 items were not related to the target structure of the study. The distractors were not actually counted in the scoring of the test. On this test, each correctly answered item would receive one point, thus, the total possible scores of the test was 30.

The second test administered to the students during the pre-assessment session was the EIT. The EIT, which aimed to gauge learners' automated knowledge of the target structure, required the students to listen to the recording of a native speaker of English reading 35 items with an interval of 8 seconds in between each of the items. The students had to decide during that interval if the items they listened to were true about their life or not, and then repeat the item in a grammatically correct fashion. The students would receive 1 point for repeating the items in a grammatically correct way. The total score that student could obtain in this test was 30 as there were five items used as distractors which were not counted in scoring of the test. Having done the EIT which took approximately 6 minutes, the students of the TBLT group were administered the Task Assessment.

The Task Assessment was a robbery task which required students to read a hypothetical story about a robbery which occurred in Miami, then rewrite the story in the past passive voice by changing the original active voices used in the story, and finally add

an ending to the story based upon their own imagination (Appendix C). The active structures in the story were underlined in order to lower the task complexity and changing the task into a focused task. At the end of the task, each student read their ending to the story and finally one of the students ending to the story was chosen by votes of the whole class to be the best ending. The total number of sentences to be changed in the story was 17 and the students received one point for successfully changing the past active voice into the passive voice.

The PPP group included 12 students. Their instructor first explained to them how to do the GJT. It took the class around 20 minutes on average to take this test. Afterwards, the class was administered the EIT, which they had to listen to the recording and then answer each item within 8 seconds. This test caused a little bit of cacophony. As a result, the instructor divided the class into groups of four to take the test one group after the other. The Task Assessment was the last test the students of this class were administered where they rewrote the task by adding an ending to it. There were nine students in the control group. The same procedures went on for the Control group in which there were nine students. The students of the Control group also had a class on Tuesday and Thursday. They were first administered the GJT, then the EIT, and finally Task Assessment.

4.5.3. The PPP Treatment

As PPP approaches consist of the three phases of Present, Practice, and Produce, so did the PPP treatment of the study. In the first phase, i.e., the Present phase, the instructor spent the first 20 minutes of the class explaining the English past passive voice. In doing so, the instructor explained and reviewed the simple past active voice first, and then went on to explain the past passive voice. In this phase, the instructor presented the class with a

table of regular and irregular verbs in English in their past and past participle form. The instructor made sure that students understood the concept of the past participle form of verbs. The students were presented with ample examples of how to change an active voice to a passive voice and vice versa.

After the present phase, the instructor moved on to the practice phase where students practice what they have just learned. In this phase, which took about 30 minutes, the instructor presented the class with three drills. As for the first drill which included ten items, the students had to complete each item through filling in the blanks by making the appropriate changes to the simple forms of the verbs in the parentheses so that each sentence would be changed into a past passive one. The second drill in this phase was a cloze passage with seven blanks where students had to choose between two forms of the verb: the active voice and the passive voice. In order to do so, students needed to read the cloze passage and comprehend it and then choose whether the active or the passive form of the verb would best complete the whole passage. As for the third drill, the students had to change the items of the test from the active voice to the passive voice. In front of each item of the test, the drill provided the students with the right word order of the passive voice in the parentheses for each of the items which were to be changed to. For instance, one of the items of the test was as follows: “We sold tickets for all shows at the Box Office,” which students had to change to the passive voice. In front of this sentence, the right word order of the subject, verb, and object for the past passive tense was presented such as “Tickets for all shows/sell/at the Box Office.” It should be noted that only the right word order was indicated in the parentheses and the verb was in its simple form so as the students would have to make the necessary changes to it in order to make a past passive sentence.

In the produce phase of the PPP treatment, the instructor made use of four oral drills of repetition, substitution, restoration, and backwards build-up. Each of these oral drills included ten items. As for the repetition drill, the students had to simply repeat the sentences in the English past passive voice mode after the teacher. The instructor broke each of the sentences into smaller chunks during the repetition drills. In terms of the substitution drill, the instructor had the students repeat the sentences after him by changing the sentences according to the prompt given to them. As an instance, the teacher read the following sentence to the students: “The cats were fed by Sarah” and then provided the students with the following word as a prompt: “they.” The students changed the sentence into “They were fed by Sarah.” In addition, the restoration drill included items with three words that students had to make a past passive sentence with. The instructor read the three words to the class and the students repeated it in a past passive voice. Finally, in the backwards build-up drill, as the name of the drill indicates, the instructor started the drill by breaking the sentences of each item into three chunks. The instructor then repeated the chunks from the end of the sentence towards the beginning until the sentence was repeated in its complete form by the students. As an example, the instructor first read out loud the phrase “by my aunt” and had the class repeat it accompanied by individual spot-checks, then the teacher added two words making the phrase longer by reading “was prepared by the aunt” again having the students repeat it accompanied by individual spot-checks. Finally, the teacher read the whole sentence of “the tea was prepared by the aunt,” which was a past passive sentence together with spot-checking some single students.

4.5.4. The TBLT Treatment

The TBLT treatment included three instructional tasks. The first task, which was a picture task, required the students to help a hypothetical character in the task named Rebecca write a letter to her brother, who had moved away, about the changes in the household which transpired ever since he left. There was a total of nine changes depicted through pictures. There were two pictures for each of the nine items, one belonging to the time before Rebecca's brother left and the other for after the time her brother left. There was also a verb used as a prompt on top of an arrow pointing towards the new picture indicating the change. The students had to look at the picture and by using the prompt verb make a sentence in the past passive voice explaining one of the changes that occurred over when Rebecca's brother was away. The students had to write their sentence on the lines below the two pictures of each item.

The instructor used Willis and Willis (2007) model for implementing the tasks. Willis's model includes the following phases: a pre-task, a task cycle, and a language focus. In the pre-task phase, the instructor should activate students' background knowledge and warm their minds up for performing the task. This could be done by asking questions relevant to the topic of the task. Additionally, the teacher should make sure that students understood the instructions of the task by explaining it thoroughly. In the task cycle, the students first do the task in pairs or groups in what is called the task stage, then the students should get prepared to report to the class either orally or in a written mode how they went about doing the task and how they planned to undertake the task. This is called the planning stage of the task cycle. Afterwards, in the third phase of task the task cycle which is called the report stage, one or some pairs or groups are selected to actually report to the whole

class about how they planned the task. The language focus phase of Willis and Willis's model included two stages namely, analysis and practice. In the analysis stage, the examine and talk about the language features in the task that they deemed interesting and important, and finally in the practice stage, the instructor dwells on the linguistic forms by reviewing words, grammatical items and patterns in the task trying to direct students' attention to the intended linguistic features of the task (Willis & Willis, 2007).

In conjunction with Willis and Willis's model, the instructor in the present study started the pre-task phase by asking questions in order to activate students' background knowledge such as what changes did each of the students' household have over the past few years? Then he went on to explain the instructions of the task making sure that everyone grasped it. Afterwards, in the task cycle phase, the instructor broke the class into pairs having them perform the task, then some pairs reported to the class how they did the task and compared their results. Having done all of these phases, the teacher introduced the language focus phase where he started giving instruction about the grammatical point of the task, i.e., the English past passive voice, followed by practicing the structure through going over the similar examples that students could give about the changes that they have done in their life.

The second treatment task was a focused task that required students to complete a text message that Cindy, a character in the task, was to send in reply to her mom's note. The task scenario was about Cindy who received a note on the kitchen table from her mom asking her to do some chores for her while she was away for a couple of days. Then Cindy decided to send her mom a text ensuring her that she did what she was supposed to. What the task required students to do was to help Cindy finish her text using the English past

passive voice. The text was an incomplete note including ten blanks that students had to fill in with the past passive voice.

The instructor started implementing the task by asking if any of the students ever had to take care of the house when their parents were away, and if *yes*, was there any interesting or funny story to share with class? Then the instructor wrapped up the pre-task phase by making sure that the students understood the instructions of the task. Having implemented that, the instructor had students work in pairs performing the task. Like the first task, in this phase of task cycle, after the students performed the tasks in pairs, they needed to discuss how they approached completing the task and then compare their results with their classmates. At the end of the task implementation, the instructor went on giving some analysis of the target structure and practicing some examples in the task with students. The instructor tried to highlight the copula and past participle part of the past passive voice by referring to some sentences in the task.

The third treatment task followed a less direct way of drawing the students' attention to the target structure. The task included two parts: a) a reading passage where the students needed to read John's father explain to him about the process of traditionally making wheat bread, and b) an activity requiring the students to put the sentences in the same chronological order that John's father explained about the process of making wheat bread. In fact, the best way to describe a process is to use the past passive voice as the doer of the action is not important but what happens in the process is significant. Therefore, this task used the past passive voice in describing the process of making bread and in order for students to be able to successfully perform the task, they had to have a knowledge of the past passive voice to decode the sentences.

The instructor started the task by giving some warm-up by asking questions of the students if they ever thought how bread was made. He also asked if anyone works in a bakery or knew how to bake bread or cake. Having asked the warm-up questions, the instructor explained the instructions of the task making sure that everything was clear to students as to how perform the task. Then he asked the students to work in pairs and do the task. The instructor also asked the students to report orally to the class how they did the task and then compare their results with other pairs in the class. After the task implementation, the instructor gave students a language focus by explaining the target structure examples in the task.

4.5.5. The Control Group Treatment

As for the control group activities, the instructor worked on five reading passages adopted from *Steps to Understanding* (Hill, 1980). The passages, which narrated funny stories, were at the intermediate level according to the book. As for the first passage, which was the story of Peter in the army, the instructor had a discussion about 20 minutes with students where he asked questions about whether there should be a compulsory military service in Iran or not, which was somehow more of a tangible subject for students to talk about as most of them had to deal with it. Afterwards, the instructor set about reading the passages and explaining the difficult words and structures. In the end, the students had to answer a comprehension question, as well as a set of grammar questions about the difference between the object pronouns and reflexive pronouns. The second passage narrated the story of Mr. Richards and problems regarding guests in his summer resort. The instructor led a discussion about how the students spent their summer vacation and what their plan would be for next summer for about 10 minutes. Then he went about reading and explaining the

text. The text was followed by a true-false, an open-ended set of comprehension questions, and a short cloze passage that students had to fill out. The third passage recounted the story of Helen and the cop. The instructor started this passage by asking students if they ever had a run-in with the law or received a ticket, and if *yes*, they were asked to share their experience with the class. Afterwards, the instructor read and explained the passage. At the end of the third passage, there were true-false, an open-ended comprehension questions, and a picture-sentence matching activities. The fourth passage was about a comic story of Mr. Thompson and his son at the bar. The instructor asked students warm-up questions about their ideas about drinking as well as their relationship with their father. Upon finishing the warm-up, the instructor read the passage and explained ordinal number in the meantime. Having finished reading the passage, the students answered a set of true-false questions as well as a set of comprehension questions. The fourth passage also included a picture activity where students had to match some sentences about the passage with their corresponding pictures. Last but not least, the fifth passage narrated the story of Jim and the thief. Like other passages the instructor gave a warm-up of 10 minutes by asking students if they have ever been robbed or witnessed a robbery. In the same fashion as other passages, the instructor then started reading and explaining the passage followed by a true-false, an open-ended comprehension questions, and synonym-antonym activities.

4.5.6. Post-assessment

The post-assessment followed the treatment session for each of the groups. All three groups of TBLT, PPP, and Control had to take all four sections of the post-assessment in the same fashion as they did for the pre-assessment. The post-assessment tests were the same as the pre-assessment tests with the only difference being that the order of the items

for the tests of the GJT and EIT have been changed for the post-assessment in order to prevent to the practice effect of the tests on the students' performance.

In the post-assessment session, the instructor administered the GJT first where the students were required to judge the grammaticality of 40 items and, in case of finding errors, they had to write the correct form thereof underneath each item. The instructor explained the instructions for the GJT and made sure that everyone understood how to approach the task. Having administered the GJT, the instructor asked the students of each group to take the EIT. The EIT test took about 5 minutes as the students had 8 seconds to answer and repeat 35 items. Afterwards, the narrative task was administered to each group. The instructor made sure that students well understood the instructions of the tasks by explaining the tasks and asking warm-up questions. As for the first narrative task, the students had to read a passage about a fabricated robbery which occurred in Miami, then they had to change the active sentences which were underlined to passive sentences. Overall, like the pre-assessment test, there were 21 sentences to be changed in the passage.

The pilot studies of this research have been conducted by using both a pool of native speakers of English as well as a pool of English language learners. In order to check the validity and reliability of the assessment tool, first the researcher had 18 native speakers of English who were undergraduate students of Education at Florida international university take the Task Assessment. The students took the Task Assessment with no time limitation. The problem that arose in the beginning of the test was that the students had trouble understanding the instructions of the Task Assessment due to the presence of the linguistically technical term *the passive voice*. The students kept asking the researcher what the meaning of *the passive voice* was and how they were supposed to approach doing the

tasks. The researcher presented them with some examples of the passive and active voice until everyone was clear what the test asked them to do. Aside from this problem, the students did not seem to have any other issues with the test and answered the test with perfect accuracy.

CHAPTER 5. RESULTS

Overview

This chapter first examines the normality of data by running a Shapiro-Wilk test to find out what statistical test would better fit the analysis. The results of the Shapiro-Wilk test indicated that the data was not normal; therefore, a set of non-parametric tests were used to further analyze the data. Along the same line, a Kruskal-Wallis test was conducted so as to evaluate the possible difference among the means of the three groups in the pre-assessment stage before the treatment was given to the students (Hollander & Wolfe, 1999). The Kruskal-Wallis test was run to ensure that the difference among the means of the three groups was negligible enough to attribute the possible differential effects in post-assessment to the treatment of the study. The results of the Kruskal-Wallis test failed to show any differences. Additionally, a set of Wilcoxon Signed Rank tests were applied for each of the tests to pinpoint any within group differences (Hollander & Wolfe, 1999). The Wilcoxon Signed Rank test indicated that the TBLT and PPP group outperformed the Control group, and TBLT group performed better than the PPP group in the Task Assessment. Finally, a Kruskal-Wallis test was run for the post-assessment which indicated no between-group differences after the treatment.

As for the first research question which inquired about the effects of the PPP and TBLT treatments on the students' performance on the GJT, the results of the study indicated that the PPP and TBLT group's performance significantly improved as compared to the Control group. As for the second research question which investigated the effects of the PPP and TBLT treatments on the students' performance of the EIT, the study yielded similar results where the PPP and TBLT groups outperformed the Control group'

performance. The third research question addressed the effects of the PPP and TBLT treatment on the students' task performance. The results indicated that only the TBLT group managed to significantly improve their performance. The PPP and Control groups failed to obtain significant improvement of Task Assessment.

5.1. Test of Normality

A test of Shapiro-Wilk was run in order to assess the normality of the data. If the data were normal, then a parametric test would be used in order to analyze the data. In case the data failed to be normal, a non-parametric test would be used for the data analysis. It should be noted that the Shapiro-Wilk test works best when the number of participants is fewer than 50. The results of Shapiro-Wilk test revealed that the data for GJT 1, EIT 2, and Task Assessment were not normal ($p < .05$). Below Table 3 shows which group did not have a normal data on which test.

Table 3. Shapiro-Wilk Test of Normality

<i>Test</i>	<i>Group</i>	<i>Statistic</i>	<i>df</i>	<i>p</i>
GJT 1	TBLT	.86	13	.04
EIT 2	TBLT	.85	13	.03
Task 2	TBLT	.82	13	.01

5.2. Pre-assessment Analysis

Given that the data did not meet the requirement of normality, a Kruskal-Wallis test was conducted in order to evaluate the difference among the means of groups in the pre-assessment phase before the treatment was given to the students. The test of Kruskal-Wallis indicated that there was not any significant difference among the three groups of TBLT, PPP, and Control on any of the tests of the GJT, the EIT, and Test task. In fact, the

test of Kruskal-Wallis showed that the groups of students had similar performance on the three tests, showing to be at a roughly similar performance level. In other words, there was not a significant between-subject difference among the three groups of TBLT, PPP, and Control in the pre-assessment phase of the study ($p > .05$). The results of the Kruskal-Wallis test are shown below in Table 4.

Table 4. Kruskal-Wallis Test for Pre-Assessment Analysis

Test	Group	N	Mean	Std. Deviation	Chi-Square	<i>P</i>
GJT1	TBLT	13	17.30	6.60	3.405	0.182
	PPP	12	20.58	4.07		
	Control	9	22.33	3.93		
EIT1	TBLT	13	17.30	5.08	0.316	0.854
	PPP	12	16.50	4.01		
	Control	9	17.33	2.87		
Task11	TBLT	13	13.42	2.00	0.767	0.682
	PPP	12	12.87	2.16		
	Control	9	13.61	1.83		

Owing to the fact that the data were not normal, a Wilcoxon Signed Rank Test was run in order to see the differential performance of each of the groups on each pair of the tests of the GJT, the EIT, and Task Assessment. This test would allow to see which group performed better on a single test in comparison with other tests which is examined in the following section.

5.3. Non-parametric Analysis

5.3.1. The Grammaticality Judgement Test

The mean score for the GJT as well as the standard deviation of all the groups of TBLT, PPP, and Control group on pretest and posttest for this test are displayed in Table 5. As shown in the table, the mean of each group's performance increased ranging from

17.30 to 22.33 on pretest to 21.07 to 22.39 on posttest. The mean scores in the table indicates that TBLT group followed by the PPP group had the highest increase in the means. The standard deviation of the GJT scores of the groups shows more variation in the score on pretest ranging from 4.02 to 6.60 than posttest ranging from 9.83 to 10.78.

Table 5. Means and Standard Deviations for the Grammaticality Judgement Test

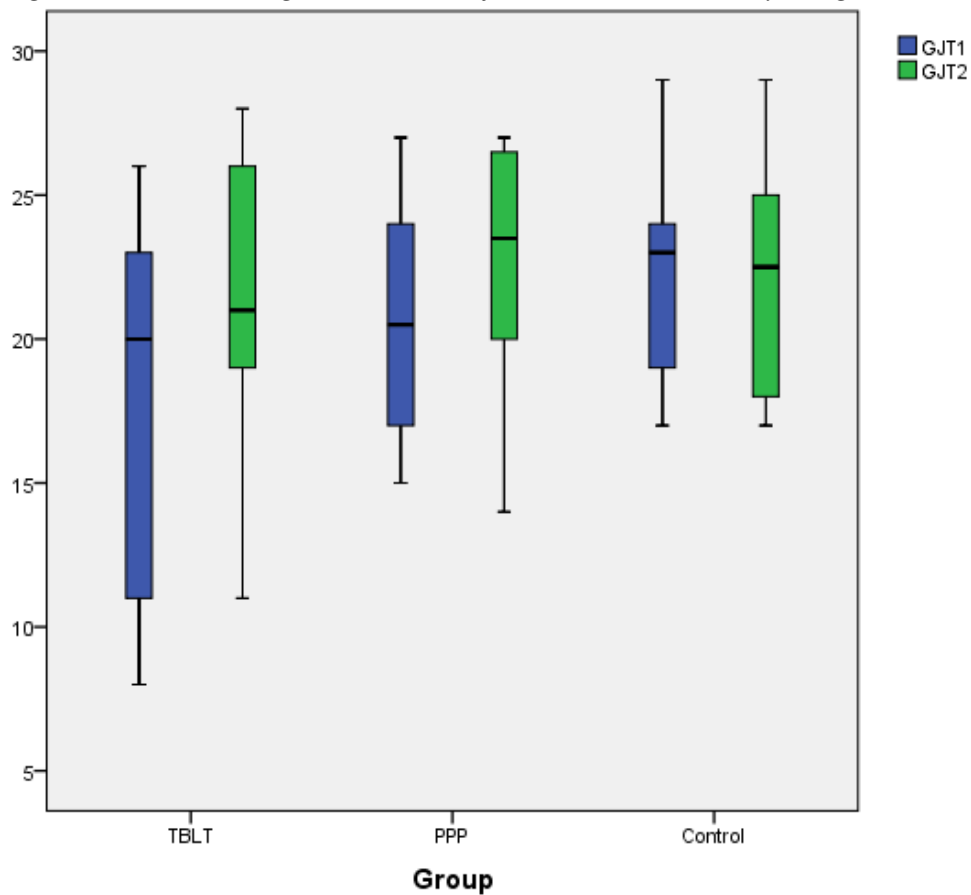
GJT	Pre-test			Post-test		
	n	M	SD	n	M	SD
TBLT	13	17.30	6.60	13	21.07	5.21
PPP	12	20.58	4.07	12	22.58	4.37
Control	9	22.33	3.93	9	22.39	4.06

A Wilcoxon Signed Rank Test was conducted to examine the differential effects of the treatments of TBLT, PPP, and Control conditions on students' use of the past passive tense in English. The results of the Wilcoxon Signed Rank Test revealed a significant difference for the performance of TBLT group from the pretest ($M = 17.30$, $SD = 6.60$) to the posttest ($M = 21.07$, $SD = 5.21$), $Z = -2.52$, $p < 0.05$. The effects size value suggests a moderate practical significance for this group on the EIT, $d = -.50$. The mean of the ranks in favor of the pretest was 3.50 and the mean of the ranks in favor of posttest was 7.50. Additionally, the Wilcoxon Signed Rank Test confirmed a significant effect for the PPP group on the students' use of the target structure from the pretest ($M = 20.58$, $SD = 4.07$) to posttest ($M = 22.58$, $SD = 4.37$), $Z = -2.15$, $p < 0.05$. The effect size value for this group, like TBLT group, indicates a moderately practical significance, $d = .51$. The mean of ranks in favor of the pretest was 2.25 and the mean ranks in favor of the posttest was 5.80.

As for the control group, the Wilcoxon Signed Rank Test failed to show any significant effect for the control group performance between the pretest ($M = 22.33$, $SD =$

3.93) and the posttest ($M = 22.39, SD = 4.06$), $z = -0.73, p > 0.05$. The effects size value on the EIT for the control group suggest a small practical significance, $d = -.15$. Furthermore, the mean ranks in favor of pre-test in this group was 4.38 while the mean ranks for the post test was 4.62. The results of Wilcoxon Signed Rank test for all of these groups are shown below in Table 6. Figure 2 depicts the results of the Wilcoxon Signed Rank Test for the GJT.

Figure 2. Wilcoxon Signed Rank Test for the Grammaticality Judgement Test



A post hoc power analysis was conducted using the software package, G Power. The sample size of the TBLT group was 13 for the statistical power analyses and the effect size used for this assessment was 0.5. The alpha level used for this analysis was $p < .05$. The post hoc analyses revealed the statistical power for this study was relatively small at 0.50. Additionally, the statistical power for the PPP group with effect size of .51 and Control group with effect size of .017 was .47 and .05 respectively, as shown below in Table 6.

Table 6. Wilcoxon Signed Rank Test for Grammaticality Judgement Test

<i>Groups</i>	<i>N</i>	<i>r</i>	<i>z</i>	<i>p</i>
TBLT	13	.50	-2.52	.012
PPP	12	.51	-2.15	.031
Control	9	.017	-0.73	.94

In sum, the results of Wilcoxon Signed Rank test showed that the performance of the TBLT and PPP group significantly improved on from pre-assessment to post-assessment while the control group failed to achieve significant results.

5.3.2. The Elicited Imitation Test

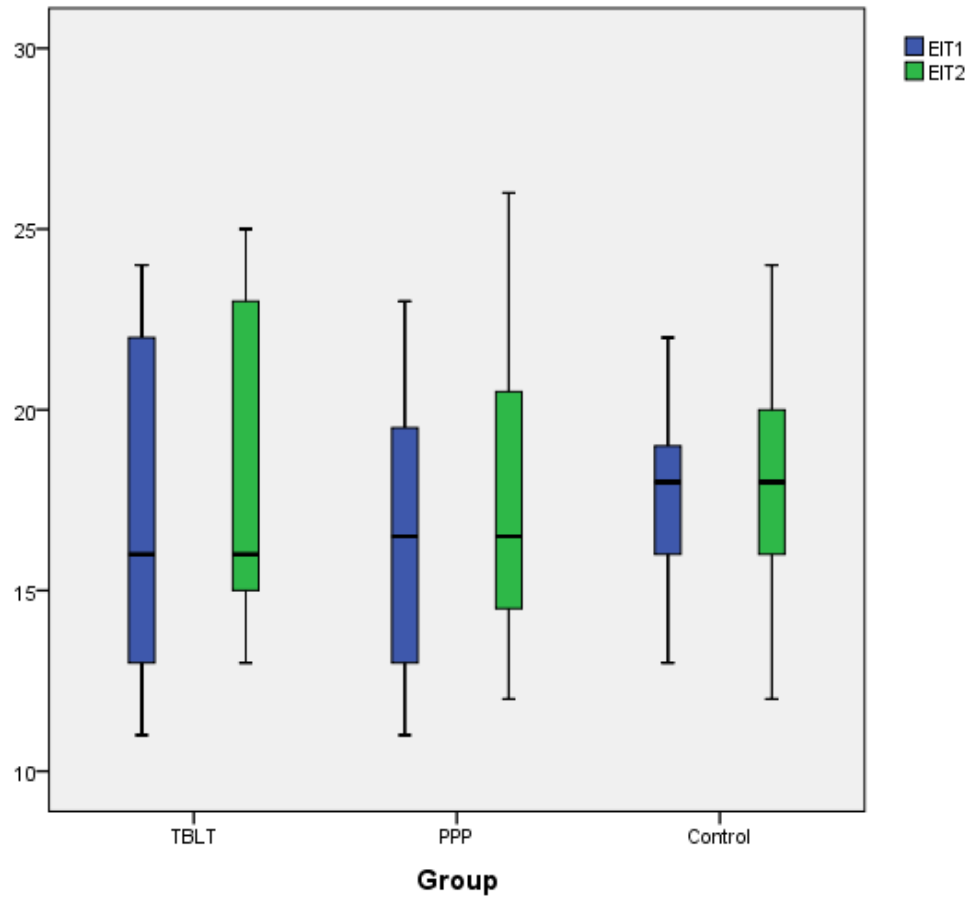
Table 7 below provides the descriptive statistics for the EIT. The mean of each group's performance increased, ranging from 16.75 to 17.30 on pretest to 17.33 to 18.61. Additionally, the standard deviation for the groups showed more variation for the two groups of PPP and Control ranging from 2.87 to 4.01 on pretest to 3.44 to 4.16 on posttest. However, the TBLT group had less variation from pretest to posttest ranging from 5.08 on pretest to 4.55 on posttest.

Table 7. Means and Standard Deviations for Elicited Imitation Test

EIT	Pre-test			Post-test		
	n	M	SD	n	M	SD
TBLT	13	17.30	5.08	13	18.61	4.55
PPP	12	16.50	4.01	12	17.50	4.16
Control	9	17.33	2.87	9	17.88	3.44

The Wilcoxon Signed Rank Test was run for the EIT assessment and the results showed a significant effect for the TBLT and PPP group but not for the control group. The Wilcoxon Signed Rank Test confirmed a significant effect for the TBLT group on student's performance of the EIT, $Z = -2.58, p < 0.05$. The mean of the ranks in favor of the pretest was 2.50 and the mean of ranks in favor of posttest is 5.83. The effect size value on the Task Assessment for this group indicated a moderate practical significance, $d = .54$. Along the same line, the Wilcoxon Signed Rank Test showed a significant result for the PPP group on the students' use of the target structure, $Z = -2.20, p < 0.05$. The mean of the ranks in favor of pretest was 2.50, while the mean of the ranks in favor of the posttest was 4.80. Furthermore, the effect size value on the Task Assessment for the PPP group suggested a relatively moderate practical significance, $d = .45$. Nonetheless, the results of the Wilcoxon Signed Rank Test did not yield any significance for the Control group, $z = -.660, p > 0.05$. Additionally, the mean of the ranks in favor of the pretest was 4.25, and the mean of the ranks in favor of the posttest was 5.60. The effect size value on the Task Assessment for the Control group showed a very small practical significance, $d = .015$. The results of the Wilcoxon Signed Rank Test are shown in Table 8. The results of the Wilcoxon Signed Rank Test for the EIT is depicted in Figure 3.

Figure 3. Wilcoxon Signed Rank Test for the Elicited Imitation Test



The post hoc power analysis revealed that the statistical power for the TBLT, PPP, and Control group was .55, .41, and .05 respectively.

Table 8. Wilcoxon Signed Rank Test for the Elicited Imitation Test

Groups	<i>N</i>	<i>r</i>	<i>Z</i>	<i>p</i>
TBLT	13	.54	-2.58	.010
PPP	12	.45	-2.20	.028
Control	9	.015	-0.66	.509

In summary, the results of the Wilcoxon Signed Rank test indicated that the TBLT group outperformed the PPP and control group in the EIT. The Control group failed to achieve any significant results.

5.3.3. Task Assessment

The descriptive statistics of each group's performance on the Task are shown below in Table 9. The descriptive statistics for the Task Assessment indicated that there was an improvement for the mean of all groups ranging from 12.87 to 13.61 on pretest to 13.58 to 15.20 on posttest. Additionally, the standard deviation of the groups over the course of time shows that there is an overall less variation for all groups from pretest ranging from 1.83 to 2.16 to posttest ranging from 1.60 to 1.99 on the posttest.

Table 9. Means and Standard Deviations for Task Assessment

GJT	Pre-test			Post-test		
	n	M	SD	n	M	SD
TBLT	13	13.42	2.00	13	15.20	1.60
PPP	12	12.87	2.16	12	13.58	1.99
Control	9	13.61	1.83	9	14.16	1.80

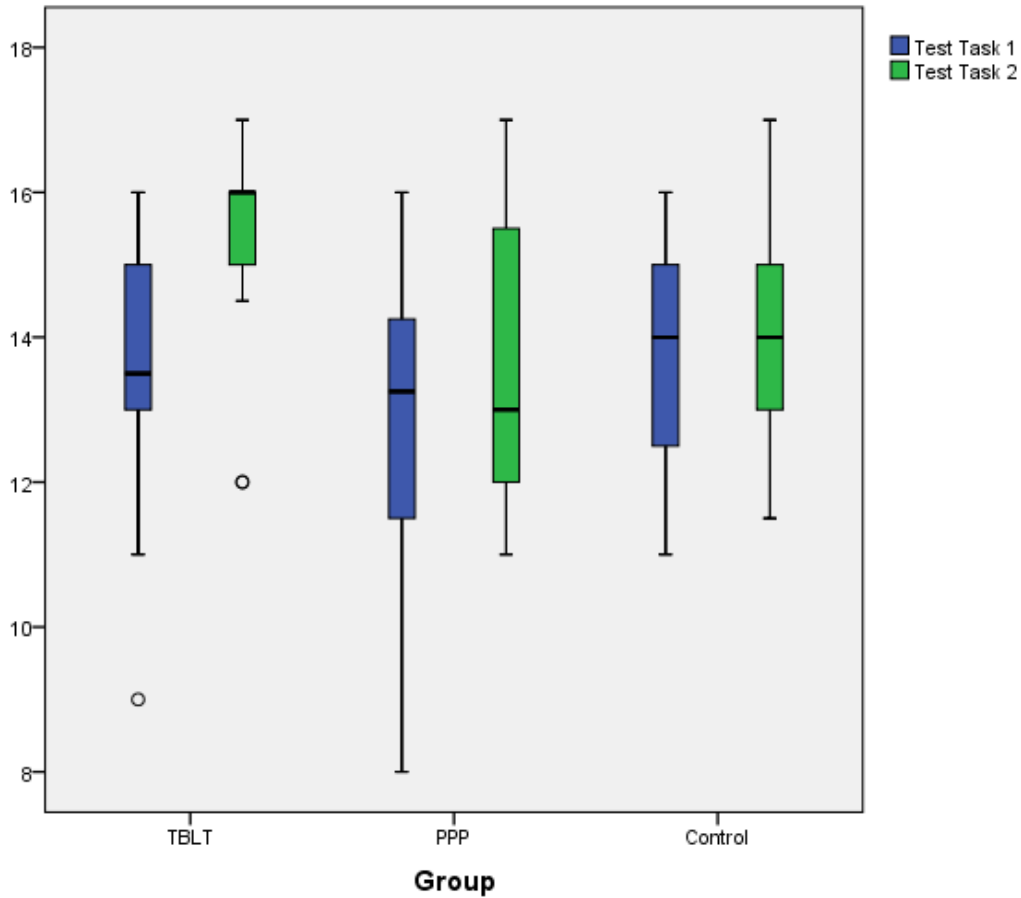
The Wilcoxon Signed Rank Test indicated that there were only significant results for the TBLT group on the Task Assessment ($M = 13.42$, $SD = 2.00$), $z = -2.76$, $p < 0.05$, while there was no significant results for the PPP and Control groups. The mean of the ranks in favor of the pretest was 2.00 and the mean of ranks in favor of posttest is 7.40. The effect size value on the Task Assessment for the TBLT group suggested a moderate practical significance, $d = .54$. The Wilcoxon Signed Rank Test for the PPP group revealed that their performance narrowly missed the level of significance ($M = 12.87$, $SD = 2.16$), $z = -1.93$, $p > 0.05$. Moreover, the mean of the ranks in favor of the pretest was 3.83 and the mean of ranks in favor of posttest is 6.81. The effect size value on the Task Assessment for the PPP group indicated a small to moderate practical significance, $d = .39$. As for the Control group, the Wilcoxon Signed Rank Test indicated no significant difference between

the performance of the students in this group on two phases of pretest and posttest ($M = 12.87$, $SD = 2.16$), $z = -1.34$, $p > 0.05$. The mean of the ranks in favor of the pretest was 4.25 and the mean of ranks in favor of posttest is 4.58. The results of Wilcoxon Signed Rank Test are shown in Table 10 and figure 4. Additionally, the effect size value on for this group showed a small practical significance, $d = .31$. The post hoc power analysis revealed that the statistical power for the TBLT, PPP, and Control group was .55, .34, and .22 respectively.

Table 10. Wilcoxon Signed Rank Test for Task Assessment

<i>Groups</i>	<i>N</i>	<i>r</i>	<i>z</i>	<i>p</i>
TBLT	13	.54	-2.76	.006
PPP	12	.39	-1.93	.054
Control	9	.31	-1.34	.180

Figure 4. Wilcoxon Signed Rank Test for Task Assessment



5.4. Post-Assessment Analysis

In the end, in order to see if the treatment was effective enough to have resulted any difference among the three groups, a Kruskal-Wallis test was conducted one more time. In other words, as opposed to the Wilcoxon Signed Rank Test which was assessing the within-group differences, the Kruskal-Wallis aimed at finding between-group differences. The test of Kruskal-Wallis indicated that there was no significant difference among the three groups of TBLT, PPP, and Control on any of the tests of the GJT, the EIT, and Task Assessment ($p > .05$). The results of the Kruskal-Wallis test are shown below.

Table 11. *Kruskal-Wallis Test for Post-Assessment Analysis*

Test	Group	N	Mean	Std. Deviation	Chi-Square	<i>P</i>
GJT1	TBLT	13	21.07	5.21	0.52	0.76
	PPP	12	22.58	4.37		
	Control	9	22.38	4.06		
EIT1	TBLT	13	21.95	4.55	0.39	0.82
	PPP	12	18.61	4.55		
	Control	9	15.19	4.16		
Task11	TBLT	13	17.88	3.44	4.40	0.11
	PPP	12	18.02	4.05		
	Control	9	13.61	1.60		

Therefore, based on the data analysis, we can conclude that the treatment was effective, but only to the extent that there was a within-group improvement. This treatment, however, was not effective to the extent it could cause any between-group difference. In fact, the Wilcoxon Signed Rank Test revealed that the treatment was effective in helping the TBLT group on all the tests and PPP group on only two tests of the GJT and EIT.

CHAPTER 6. DISCUSSION

Overview

The study was designed to compare the differential effects of TBLT and PPP treatment on language learners' performance on the GJT, the EIT, and Task Assessment. The results of the study indicated that the TBLT and PPP groups have had significant performance improvements on the GJT and EIT, while the Control group did not have any significant performance improvements on any assessments. The results also showed that the TBLT group was the only group that had a significant performance improvement on the Task Assessment. The present study was innovative in that it used a real task to measure learning in a more comprehensive way rather than just using non-task measures to examine the effects of even TBLT. This was actually the gap in the literature that the study originally aimed to address, i.e., the lack of the use of Task-based Language Assessment.

Another strength of the study was the use of real teachers who were the participants' real teacher in the study. In fact, the study used actual classroom teachers rather than an external teacher or the researcher as the teacher so as to enhance the authenticity of the classroom researcher and ultimately the results obtained. This chapter will now address each assessment type and discuss their results theoretically and in reference to some theories of Second Language Acquisition. The chapter first addresses the results of the GJT and Explicit knowledge which are related to the first research question. Then the results of the EIT and automated knowledge that are related to the second research question are discussed. Finally, the results on task assessment and task performance are examined. The chapter also takes into account Task-based Language Assessment, the role of the teachers, and finally the Task-based contextual knowledge.

6.1. The Grammaticality Judgement Test and Explicit Knowledge

The results of the Wilcoxon Signed Rank Test on the GJT revealed that the TBLT and PPP treatments had significant effects on their respective students' performance and the mastery of the past passive structure in English, while the Control group's performance did not yield any significant results on the mastery of the target structure. Even though the TBLT group did not have explicit instruction like the PPP group did in their treatment, they managed to have a more significant improvement than the PPP group ($M = 17.30$, $SD = 6.60$) to the posttest ($M = 21.07$, $SD = 5.21$), $Z = -2.52$, $p < 0.05$. As for the PPP instruction, the students in this group were presented with focus on forms instruction. That is, they were presented with the instruction on the past passive voice with no inclusion of communicative context. In other words, the PPP group practiced the target structure with no reference to the real-life and authentic context of language use; they just practiced the target structure in discrete sentences while their focal attention was on the use rather than the usage of the target structure. This adds more to their knowledge about the target structure than the knowledge to utilize the target structure which helps them more in the actual communicative context. In fact, knowledge about the structure of the language helped the PPP group more in forming the explicit knowledge. Nevertheless, the TBLT instructions helped the TBLT group to have a focus on form instruction where they incidentally acquired the target structure in the context of language use where their focal attention was directed towards the meaning of the message to be conveyed than the form. In other words, the TBLT instructions can be effective for helping students master the target structure indirectly and incidentally by having them focus on the meaning conveyance of the task during their performance. Additionally, the focus on form

opportunity that the TBLT group had in the analysis phases of task implementation gave them the chance to relate the structure to the context of its use. Along the same line, the results of the study highlight the higher effects of focus on form over focus on formS in preparing students not just for the task performance but also on traditional methods of assessment such as the GJT and EIT. The interesting point of the results of this study is that the focus on form was effective even on the GJT, which tends to be a test tailored more for the focus on formS instructions. The Control group did not achieve any significant effect on their performance on posttest as they did not have any preparation for the tests on the intended structure. The Control group just had a discussion based reading which did not include any target structure.

Another perspective through which the results of the GJT could be analyzed is to consider the argument of explicit/implicit knowledge or otherwise known as declarative and automated knowledge. As explained in Chapter 1, explicit knowledge is a type of knowledge which is encyclopedic and factual in nature such as the knowledge of abstract grammatical rules (Ellis, 2005). In order to obtain it, learners should consciously be aware of or analyze the relations and rules between different elements of the target feature. On the other hand, the implicit knowledge is automated; that is, learners gain more control over the rules and fragments of language over time to the degree that little or no conscious attention is needed to apply those rules. The study hypothesized that the PPP instruction would build and improve the PPP group's explicit or declarative knowledge. In other words, previous researchers believed that through the PPP instructions, the students of the PPP group would master the grammatical rules of the target structure through focal attention to the target structure, and therefore they could outperform the TBLT group on

the test of the GJT. However, surprisingly, this study showed that the TBLT group had a more significant performance than the PPP group on the test which was intended to measure explicit or declarative knowledge. One of the arguments that could be put forth to support this result is the effect of the language focus phase of the task implementation. During the language focus phase, the teacher tended to examine, discuss, and practice the intended structure of the task, in this case being the past passive voice. It can be concluded that the language focus phase of the TBLT instruction well prepared the students of this group with the explicit knowledge that ultimately helped them perform well on the GJT. Additionally, the feedback that the instructor provided during the TBLT students' performance of the task could have had its effective influence on helping learners develop explicit knowledge of the target structure. Ellis (2005) holds that learners can even figure out grammatical rules of target structures by just attending to them and analyzing them after they have performed the task. This could also be a possible explanation for the TBLT students' superiority over the PPP students in forming explicit knowledge.

6.2. The Elicited Imitation Test and Implicit Knowledge

Similar results to that of the GJT have been achieved by the EIT where the TBLT group again had a better rate of improvement over the PPP and Control groups, respectively. The EIT, which was supposed to assess students' mastery of the automated knowledge, indicated that TBLT and PPP groups achieved significant improvements while the Control group did not obtain significant results. Considering the fact that automated knowledge is of unconscious and less controlled nature where the students do not need to attend to every specific structure to use, we could conclude that these features of implicit knowledge could better be produced through focus on form instructions. That is, in order

for students to master implicit knowledge of a particular structure, they need to be automatic in their production with little attention to the form of that structure. This is in conjunction with the nature of Task-based Instruction where students need to produce linguistic forms in the context of language use without necessarily paying attention to the rules and regulations that govern that structure. In other words, the unconscious nature of focus on the meaning to be conveyed during task performance in TBLT better prepared the TBLT students to be faster in answering the questions of the EIT, which was a timed test. On the other hand, the PPP instructions which dueled more on explaining the rule and regulation of the target structure tended to produce more mastery on the explicit knowledge of the structure as the students in the PPP group were not given the chance to use the target structure in a different context. In fact, it could be argued that the findings of the present study show that the mechanical nature of the PPP instruction could not even be as effective as uncontrolled meaning-focused TBLT instructions. In theory, it seems like the PPP group should have reached a level of automaticity to have succeeded on the EIT; however, in practice, the TBLT students were better prepared for this test due to the nature of their acquired knowledge which was less analytical and more incidental and uncontrolled. Therefore, the results of the EIT showed that TBLT instructions can help students reach the level of automaticity faster than the PPP owing to the fact that they have the opportunity to use the target language in an authentic context while this opportunity was not there for PPP students and the type of mechanical drills that they have done on the treatment seems to require more time to make students reach the level of automaticity and uncontrolled use of the target language. The Control Ggroup's lack of significant improvement could be

attributed to the fact that they did not receive the relevant instructions as to how to use the target structure, therefore, they did not show any remarkable mastery of the target structure.

Along the same line, DeKeyser's (1998, 2007) skill-learning theoretical model states that implicit automated knowledge can be fostered when language learners have access to the explicit version of that linguistic form and also the chance to apply that linguistic form in the context of communication and interaction in which their attention is frequently drawn to that linguistic form. In other words, this model holds that explicit knowledge has the potential to be converted or transformed into implicit knowledge through the context of communicative language use. The results of the study seem to be in line with DeKeyser's argument, known as strong interface position. Ellis (2005) defines the strong interface position as the fact that explicit knowledge can be converted to implicit knowledge through practice, which means language learners first learn the declarative knowledge then through rehearsal and practice reach the level of automaticity and then turn it to the implicit knowledge. The outperformance of the TBLT groups in developing students' implicit/automated knowledge could be attributed to this interface position as the TBLT group which was successful in developing the students' explicit knowledge through focus on form instruction. The focus on form instruction provided them with the opportunity to rehearse and practice their explicit knowledge in the context of language use; as a result of this rehearsal, this explicit knowledge became more automatic and ultimately turned to the implicit knowledge. As such, the TBLT condition had a better improvement rate than their PPP counterpart on the EIT which was specifically designed to assess the implicit/automated knowledge. Even in the least optimistic view, the results of this study could be explained in light of weak interface position, which holds that explicit

knowledge might not transform into implicit knowledge but in fact facilitates the statistical learning processes involved in its development (Ellis, 2005; Li, et al., 2016). Even considering the weak interface position, the results of the study indicate that TBLT fosters and sets the grounds for the mastery of implicit/automated knowledge. Clearly, the findings of the present study do not support the non-interface position, which states that implicit and explicit knowledge are formed through different acquisitional mechanisms (Ellis, 2005).

Ellis (2002) argues that the development of implicit knowledge has three processes: a) noticing, where the student becomes conscious of the presence of the linguistic feature which they had ignored in their previous encounters, b) comparing, through which the learner attends to the difference between their existing linguistic repertoire to see the gap between the input and their existing grammatical repertoire, and c) integrating, where the student adds and integrates the new linguistic feature into their existing repertoire. According to Ellis (2002), focused tasks have the potential to contribute to the development of implicit knowledge through fostering noticing and comparing in learners. He further says that focused tasks may also facilitate the *integration* conditioning that the students are developmentally ready to acquire that linguistic item. This is in line with the findings of the present study; as mentioned above, focused tasks used in this study helped learners noticing, comparing, and integrating processes, which helped them perform significantly better than other groups in the development of the implicit knowledge.

In essence, it should also be noted that in order to obtain more realistic findings of the effects of the methods of TBLT and PPP on the acquisition of implicit knowledge, we need to conduct longitudinal studies with longer amount of language instruction. In fact, the shorter period of language instruction would be biased towards better effects for the explicit knowledge, however, the implicit knowledge due to its automatic and incidental nature require longer period of time in order to develop. More importantly, TBLT instruction is not originally designed to develop explicit or implicit knowledge. The nature of TBLT has the more important goal of helping learners develop the ability to perform tasks using language as a tool. Therefore, along the same line of argument, this type of proficiency, that is the task performance, needs a longer period of time to reach its full productivity. In fact, explicit knowledge might be developed through a two-hour instruction, however, the development of implicit knowledge and task-performance proficiency necessitates longer period of instruction.

6.3. Task Assessment and Task Performance

As for the test task, the results of the Wilcoxon Signed Rank Test confirmed that the TBLT group had a higher rate of improvement than the PPP group and Control group by showing a significant effect on the performance of students on the post-test. The PPP group was very close to reaching a significant result. The results go with the research hypothesis that the TBLT group would perform better on a test that requires students to use more than their linguistic competence. In fact, the Task Assessment required the students not only to focus on the linguistics features but also to go beyond the target feature by being more focused on meaning conveyance and planning how to perform the task. As such, they need to use their communicative competence rather than just resorting to their linguistic

competence. The students needed to concentrate on the strategies and plans to approach the task which would in turn take up an important part of their working memory. For instance, the fact that the students had to come up with an ending to the story would be more demanding and complex for students than modifying a linguistic form. This was in fact unknown territory for the other two groups where just resorting to linguistic competence would not be sufficient to perform the test successfully. In fact, the results of the Task Assessment could be reviewed through the rationale that the TBLT group had practice with performing the task and using their knowledge in the context of language use, which is exactly what the PPP group lacked the instructions and practice for. Even though the PPP group performed pretty well and beyond expectations on the test task, still the performance of the TBLT group was way better than the PPP group.

The fact that the students in the PPP group did not perform as well as the TBLT group on the Task Assessment could also be attributed to the fact that tasks usually create an open-ended context of language use where the students need to use the target language with some modifications for that context which is something that the PPP group was not well prepared for as their exercises were mechanical and the changes were minor. Along the same line, performance of a task required students to rely on their resources and allows them to use whatever resources they have in their disposition to plan and perform the task; however, the students in the PPP group were not trained to face this situation. The most the PPP group was trained for was to make some minor modifications in close-ended mechanical exercises; they were not trained to make the best use of their linguistic resources in a free fashion to get the task outcome done. As a result, the performance of the TBLT group

expectedly was better than the PPP group. The Control group did not have the training nor the practice with the target language to be able to perform as well as the other two groups.

The fact that the TBLT group significantly outperformed the PPP and Control groups on the Task Test could additionally be argued in light of Long's (1983, 1989) interaction hypothesis. Long holds that the interaction that task-work fosters is significant since it serves two purposes: a) it helps the input that the students are working on to be more comprehensible through the back-and-forth negotiations that students do to make meaning; and b) during the struggle to negotiate meaning, students attend to the problematic linguistic forms in the input and output, thereby helping them to do the form-meaning relations mapping and consequently improve their linguistic output. In other words, as Ellis (2003) and Long (2015) put it, the output modification is an outcome of the interaction. In other words, the most important part of the superior performance of the TBLT group compared to the other two groups of PPP and Control could be attributed to the fact that the TBLT group had the opportunity to fine-tune its linguistic repertoire through having the chance to modify it through interaction. This is in line with the tenets of output hypothesis (Swan, 1995) which states that when learners are to produce their utterance during the interaction, they get to realize the gap between what they want to say and what they actually are able to say. In a nutshell, TBLT paves the way for the interaction among students through which the students would have richer input, and modified output. This has helped the students of the TBLT group be better prepared to take the tests especially the Test Task.

6.4. Task-based Language Assessment

The present study highlights the importance of TBLA and the necessity of its inclusion in second language research and instruction. Most of the studies (e.g., De la Fuente, 2006; De ridder et al., 2007; Gonzalez-Lloret & Nielson, 2015; Lai, Zhao, & Wang, 2011; Li, Ellis & Zhu, 2016; Shintani, 2011, 2013) that have set out to investigate the differential effects of various language methodologies including the TBLT have neglected the important role of TBLA. The present study strived to indicate that without TBLA, the evaluation of students' performance would be distorted and unrealistic. In fact, the ultimate goal of every instruction is enabling the students to use and apply what they have acquired to the real life, and language instruction is no exception to that. The declarative and automated knowledge should ultimately help learners perform with the language. The presence of these types of knowledge would be meaningful if they prove to be significantly useful in the context of language use. This is in line with the tenets of performance-based testing where students' successful performance is evaluated through how they have completed the task rather than how they used the language to do the task. In other words, in performance testing, the criterion of evaluation is not separate from the tasks itself. Furthermore, two of the findings of the present study was that a) TBLT better prepares students to perform with the language they have learned, and b) TBLA provides a more holistic measure of students' ability to *use* the language and to perform a real task.

Task-based Language Assessment has been criticized based on some grounds such as its generalizability to other contexts as well as the trade-off between the task completion and language ability of the student. As far as the generalizability of TBLA is concerned or what Long (2009) called the challenge of transferability of learners' ability

to the real world, it has been argued that TBLA falls short of having clear and straightforward criteria for classifying them so as to make sure that the pedagogical tasks and real-life tasks are of the same type. In fact, the performance nature of tasks and the holistic nature of their evaluation made the job of classifying them a little difficult. This problem has been partly solved by considering three criteria of fluency, complexity and accuracy of speech in order to classify and assess tasks more objectively. Additionally, the use of focused tasks could help remove the problem by setting a middle ground between the task completion and its language use.

Another way to make up for the issue of generalizability of tasks is to use an indirect task-based performance assessment. Directness and indirectness of the tests has to do with the degree that the test performance matches with the criterion performance. In other words, in this type of assessment, the criterion performance is subdivided into more subtasks or component steps and these subtasks are assessed separately. Therefore, these tests are analytic in the design and sample performance of certain skills (Baker, 1989; Ellis, 2003; Robinson, 1996). In fact, even though the students get to perform and complete the task, but the evaluation of their performance is indirectly related to their performance. As example of this would be the task that has been used in the present study. The students had to perform a robbery task and then add an ending to the story, however, their task has been divided into subcomponents for the evaluation. Therefore, their completion of the task was part of their evaluation and is indirectly connected to the whole score that students achieved on the test task.

The use of indirect task-based assessment in the present study made it possible to compare the students' performance on a performance-referenced test tasks with their

performance on the system-referenced assessment tests such as the GJT and EIT. In fact, if in this study the direct and holistic performance assessment had been used, it would not have been possible to compare the linguistic performance of the students with discrete system-referenced tests of the GJT and EIT. The fact that TBLA could be adjusted to be more indirect and objective may be considered its strong point as supposed to the common some scholars such as Long (2009) as its weak point. As mentioned above, TBLA can obviate its problems of generalizability through either the use of focused tasks as in this study or the use of discrete-point criteria such as fluency, complexity, and accuracy.

In fact, what inspired the present study was the fact that many studies took the important role of Task-based Language Assessment for granted, in particular Li et al.'s (2016) research. In their study, Li et al. set about examining the effects of Task-based Instruction as well as PPP on learners' performance of system-referenced tests such as the GJT and EIT. Nonetheless, TBLA is mostly set to measure performance of students in the context of language use. In other words, what was the most significant methodological flaw of Li et al.'s (2016) study was to measure the effect of TBLT using tests which were not meant to measure that nor had the validity to measure what TBLT purports to improve. Each method of language teaching should be evaluated in line with the assessment thereof so that a true picture of their influence could be evident. Along the same line, another important finding of the study was TBLT not only could improve students' performance and communication ability, but it could also achieve an acceptable level towards the goal of improving what PPP instruction aimed to improve. In this regard, we could claim that TBLT offers a more comprehensive a package of instruction than the PPP instruction since it can improve task performance as well as develop automated and explicit knowledge in

learners. And this finding could not have been achieved if the present study had not used all required assessment to achieve this conclusion.

6.5. Comparison with Li et al. (2016) Study

As this research was an attempt to replicate and improve upon Li et al.'s (2016) article, this section examines the explanations for the results of both studies. In Li et al.'s (2016) article, two task groups were used, namely focus on meaning and focus on form. The FoM was a task only treatment where students had been just required to solely perform the task without any explicit instruction or feedback, while the focus on form had the opportunity to receive some feedback. The results of Li et al.'s (2016) article indicated no significant effects for focus on meaning. In fact, they called the effects of the focus on meaning very limited and only slightly better than control group. They went ahead and put forth three arguments for the limited effect of focus on meaning on learner's performance. First, they claimed the passive structure is a late acquired form and the students in the study were not developmentally ready to acquire it (Li et al., 2016). Second, they argued they were only two tasks that provided the learning opportunity for the students which according to them, only two tasks were not enough. Last, they referred to Long and Robinson's (1998) argument that because it was a task only situation, the students did not managed to activate the cognitive processes required for acquisition to occur. As for the focus on form or the task with feedback condition, the Li et al.'s (2016) study indicated that there was a short term and insignificant effect for focus on form condition mostly on the explicit knowledge. They also argued that the short-term effect of focus on form condition is due to the nature of explicit knowledge that decays more rapidly than implicit knowledge.

Nonetheless, the most important argument that goes unnoticed in Li et al.'s (2016) study is the fact that the TBLT condition is not meant to improve the explicit or implicit knowledge but to improve the language use or performance of the learners and to measure their ability to do real tasks. Along the same line, the TBLT condition's effect must be assessed using a compatible assessment tool, which is the TBLA. In effect, the tests used in Li et al.'s (2016) study were tailored to assess the effects of the GJT and EIT but TBLT; therefore, the results of coming from these types of tests could be nothing but a distorted picture of TBLT's potential effect. Attempts were made to obviate this drawback in the present study through the inclusion of TBLA, which in fact the different effects of this study with Li et al.'s (2016) study could be viewed through this perspective. More importantly, the results of this study illustrated that the TBLT condition could in fact be even effective not only in improving task performance but also explicit and implicit knowledge, or at least better than PPP and Control conditions. This result is confirmed by Li et al.'s (2016) study where they found out that the task condition can be more effective than PPP condition in improving explicit and implicit knowledge only if task is accompanied by pre-task instruction and corrective feedback.

It should also be noted that the present study investigated the effects of TBLT in the foreign language context. In fact, TBLT tends to be even more effective in the second language context. In fact, the students in the second language context seem to be more prepared to undertake the task performance as they have everyday experience of performing tasks in the context of second language. Therefore, the results of the present study should be interpreted in light of the fact that the effect of TBLT was examined in a context that does not seem to lend itself that much to this methodology of teaching.

The present study also contributes to the classroom-based research in a foreign language context in Iran as the field of Task-based Language Teaching is at its fledging stages in Iran. In fact, theoretically, Task-based Language Teaching has attracted a lot of attention recently in Iran, however, there is still a long way to pass in order to use this methodology in actual practice. More importantly, Task-based Language Assessment is even more in need of promotion in Iran. The present study plays a small role in investigating the effects of Task-based Language Teaching and its assessment in a context where task performance might not be necessarily culturally and traditionally a common and prevalent way of language teaching.

6.6. Task-based Contextual Knowledge and Discourse Choice

Ellis (2009) holds that during task performance, both the explicit and implicit knowledge are at work. He contends that by default and naturally language learners tend to rely on implicit knowledge to perform tasks; however, in the face of difficulty in performing a task, learners exploit their explicit knowledge. He points out that the function of explicit knowledge is to help learners gain self-control in situations that are linguistically complex demanding. In fact, explicit knowledge has a mediation role in task performance which comes in handy when for instance learners are having private speech to tackle a problem. While it is true that the implicit knowledge would be used with occasional help of explicit knowledge, task performance goes beyond these two types of knowledge and trains students how to use the appropriate discourse suitable to a certain context. Take the earlier-provided example of ordering pizza on the phone. Language learners might have the explicit knowledge of how to form a polite request using “Could I have...” and they might also have reached the level of automaticity in using this structure in an unconscious

fashion. However, they might not necessarily know how to get their meaning across according to the context of ordering pizza on the phone. In fact, task performance creates a context where they can acquire the appropriate way of using a structure that fits a situation. In this regard, the learners come across the sentence “Could I order a pizza for a pick up?” as an example. In fact, while implicit knowledge can be used in task performance, task performance can set the grounds for the acquisition of the discourse which students might not get the chance to see that discourse unless during the task performance related to the context of that discourse. In other words, unless the students get to experience through tasks how to perform in ordering pizza on the phone, they will not be able to use their implicit knowledge of polite requests in English. This role of tasks is highlighted more in situations where they need to decode and understand a certain sentence. Here other examples from a daily experience that an ELL might come across in a fast food restaurant: “Do you want just the sandwich or combo?” or “For here or to go?” There are numerous expressions of this kind that have nothing to do with the nature of explicit or implicit knowledge of the learners but require the contextual knowledge where a certain linguistic form is to be used. This contextual knowledge which is intertwined with the linguistic knowledge of students is the area that TBLT claims to be able to fill the gap. In fact, TBLT is more concerned with performance proficiency or what I would call *contextual knowledge*” where linguistic knowledge, be it explicit or implicit, plays a key role. The contextual knowledge includes the knowledge of the linguistic forms that fit the situation the best as well as the knowledge of rules and regulations governing that context. TBLT paves the way for the acquisition of both of them. This might somehow justify why PPP students did not manage to perform significantly better than their pre-test even though they

had access to the implicit knowledge of the linguistic form. In fact, based upon this argument, it could be stated that the PPP group did not have the contextual knowledge necessary to perform on the test tasks.

CHAPTER 7. CONCLUSION

Overview

This chapter presents a conclusion by reviewing the results of the study in light of the theoretical and pedagogical implications. Additionally, suggestions for further research is offered for the researcher interested in this field. The chapter also presents a reflection of the challenges and obstacles that the researchers faced in the process of conducting this study.

7.1. The GJT, the EIT, and Task Assessment

The present study has been done in response to Li et al.'s (2016) article where they compared students' performance under four conditions: a) task only, where students were required to perform a task without intervention by the instructor, b) explicit instruction + task, where students were given pre-task explicit instructions of the target structure prior to their task performance, c) task + corrective feedback, where students were provided with corrective feedback by the teacher during their task performance, and last but not least d) Explicit Instruction + Task + Corrective Feedback, where students had the opportunity to have both pre-task explicit instruction as well as corrective feedback. The interesting and somewhat controversial point of Li et al.'s (2016) study was that, even though they used task-based condition in their research, they did not use any assessment compatible with TBLT. That is to say, the main methodological issue of Li et al.'s (2016) study was that they simply neglected the role of TBLA. In other words, TBLT as a full-fledged methodology of second language teaching has its own methodology of assessment. Li et al.'s (2016) study only used tests typical of the PPP methodology in order to assess TBLT's effectiveness, which naturally and automatically tends to yield distorted results.

As explained before, TBLT purports to foster communicative competence in language learners, where only a part of which is the linguistic competence. Hence, in order to assess a learner's communicative competence, there must be an assessment tool comprehensive enough in terms of assessing language learner's full communicative competence rather than just their linguistic competence. With this point in mind, the tests of the GJT and EIT fell short of assessing learner's communicative competence as they are made and intended solely for the sake of assessing linguistic competence and not more. Therefore, studies along the same line must be conducted to cover this issue and include TBLA in their evaluation of the effects of TBLT. To this end, the present study set out to show that the more authentic comparison of the PPP and TBLT includes not only different treatments but also their different relative assessments. In fact, the PPP group and TBLT should be assessed across the tests of both PPP and TBLT so that there could be a more comprehensive and realistic view of the comparison of the two methodologies of language teaching. Thus, the tests of the present study included two tests of the GJT and EIT as well as the Test Task.

The results of the study indicated that overall the TBLT group showed better performance than the PPP and Control groups, respectively. The explanation for the better performance of the TBLT group was explained through some perspectives. First, TBLT group seemed to be more effective than other groups in fostering both attention to the linguistic features as well as the meaning to be conveyed. An examination of the results of the study on all the tests confirmed that TBLT was efficacious in improving grammatical accuracy of students by having them focus on form, which occurred incidentally when students' focal attention was on meaning conveyance. In addition, TBLT was effective in

helping students use the target language in conveying meaning in an authentic context. In fact, the results of the study confirmed the better effect of focus on form, promoted by TBLT, over focus on formS, promoted by the PPP approach, even on tests that measured grammatical mastery of students. While both the TBLT and PPP groups had significant improvements on the posttest, the Control group, as hypothesized by the study, did not show any significant improvements due to not having the relevant instructional treatment.

In fact, the data indicated that the performance of both the TBLT and PPP groups was significantly better in developing the explicit knowledge than the Control group. As far as the explicit knowledge is concerned, the study hypothesized that due to the controlled nature of the exercises in the PPP treatment, the PPP group would outperform the two other groups; however, the results of the study indicated that the TBLT group showed a more significant result than the PPP and Control group. One rationale for this result could be the fact that the language focus part of the TBLT treatment prepared the students of the TBLT group to outperform the other groups. In addition, the students in TBLT group enjoyed the feedback on the side of the teacher which was to a great part linguistically directed. More importantly, due to the use of focused tasks in the TBLT treatment, the students had the chance to be directed to the linguistic feature frequently in the context where their focal attention was on meaning conveyance. This constant encounter seemed to have helped them pay attention to the linguistic feature and, therefore, master it which led to their best performance compared to other groups on the test of the GJT. Nassaji and Fotos (2011) hold that when learners get conscious of the linguistic features, they tend to notice it in subsequent communicative input. Such noticing would start the restructuring of their

implicit knowledge of the linguistic feature. This may well explain the better performance of TBLT students on the EIT test of implicit knowledge.

As for the automated knowledge, again the TBLT group showed better performance than the other groups. As mentioned above, the automated knowledge leads to a quick response which is typically unconscious and automatic. The explanation for the outperformance of the TBLT group on this could be that the students in this group had already practiced and experienced using the target linguistic features in an unconscious and incidental manner in a context where their focal attention was on the meaning conveyance. This may have helped the TBLT students to form relatively an automated knowledge. While the mechanical drills that the PPP group had did not seem to have formed the type of automaticity needed for the EIT, the control group similarly had no preparation to face this type of test. Additionally, another rationale for the better performance of TBLT group than the PPP and Control could be the fact that they had already had the opportunity to use their explicit knowledge in a communicative context. This argument is in line with DeKeyser's (1998, 2007) skill-learning theoretical model which holds that explicit knowledge could be transformed into implicit knowledge if it is used constantly and in an incidental manner in a communicative context. This is exactly what the TBLT condition offered its students: a context to practice and use their already acquired explicit knowledge. Although, it should be noted that automated knowledge tends to require longer time than explicit knowledge to develop, therefore, to gain a more realistic results on automated knowledge, we should conduct more longitudinal studies.

It should be noted that the PPP approach should be given credit for being effective in improving the student's performance in two tests of the GJT and the EIT.

Therefore, the use of the PPP approach could be justified at some phases in the classroom in the language learning process alongside the TBLT approach. Maybe the combination of both approaches might produce even more effective results which could be the focus of future research.

The study also used another assessment tool that aimed at gauging the communicative competence of the students. The study hypothesized that the TBLT group would outperform other groups on TBLA tools, since, as mentioned above, the TBLT treatment offers more than mastery over the linguistic forms. In fact, the TBLT trains students to use the language for an outcome which is not linguistic. Therefore, students need to be directed not just to the linguistic form but to the meaning conveyance as well as the context in which the linguistic forms are used. More importantly, the students need to pay a special attention to the planning phase of the task performance. That is, they need to come up with a solution which work the best in implementing the task. All of these factors lie beyond the linguistic competence somewhere within the realm of communicative competence. Additionally, the context of language use in TBLT requires that students use all of their linguistic resources freely in the process of task performance, something which might be quite perplexing to the students of PPP and Control groups. As expected, the findings of the study confirmed that the TBLT group's performance significantly improved on the posttest while the other two groups did not show any significant improvements in their results. This corroborated the fact that the PPP treatment would in the best condition to improve the linguistic competence of the students, while in the case of this study, the TBLT groups even had better improvement than their PPP counterparts due to the

flexibility of the TBLT that could help students in improving their grammatical proficiency.

The study also discovered that explicit knowledge and implicit knowledge could well be improved by both the PPP and TBLT instructions. TBLT showed that it has the potential to help learners with the linguistic form even better than the PPP condition. However, as for the third part of the research question, the findings of the study revealed that the PPP condition fell short of preparing students to be able to perform a task using the same linguistic features that they have mastered. In fact, this is a significant outcome of the study to indicate that the mere command of a certain linguistic feature does not necessarily guarantee the successful use of that linguistic feature in the context of communication. This finding might not have been obtained if the study did not use TBLA.

The study also concluded that tasks can set the grounds for the learner to learn the contextual knowledge needed to use the target feature in. As discussed above, in addition to the fact that TBLT can foster the explicit knowledge and implicit knowledge, it can provide the students with the necessary discourse that they need to master to use the target feature in the context of language use. In fact, what the learners require the most in order to communicate successfully is to know how a particular structure is used in a certain context. For example, as mentioned above, students might have mastered a certain structure and have the ability to use it automatically; however, they might not be able to use it in an automatic fashion due to the lack of contextual knowledge. In a nutshell, I would call contextual knowledge a task-based counterpart of explicit and implicit knowledge that shows students how to use a structure in conjunction with a situation and the discourse thereof.

Lastly, one of the most significant findings of the study is that both focus form and focus on formS (Long, 2001) promoted respectively by the TBLT and PPP approaches seem to be effective in developing the explicit knowledge. However, the results of the study indicate that focus on form instruction seems to be more effective in developing the implicit knowledge of the grammatical feature. This might in turn be an indicative of the superiority of TBLT over the PPP approach in that it promotes both the implicit and explicit knowledge while the PPP approach seems to be more effective with the explicit knowledge.

7.2. Task-based Language Assessment

The most important aspect of this study was to pinpoint the importance of TBLA. TBLA has been inspired by the concept of performance assessment. In other words, TBLA has introduced the broader concept of performance assessment to the field of Second Language Acquisition. In fact, what has long been neglected in the realm of Second Language Pedagogy was the ability in students to communicate in an authentic context of language use. In other words, a language learner might be quite accurate in the grammar of a second language but not able to effectively communicate in the context of language application, and vice versa, a language learner could have a broken language proficiency but be able to effectively convey their meaning. What TBLT emphasizes is attention to the context where the language would ultimately be used, both in instruction and assessment. What counts in learning a second language is not just their linguistic form, which in fact does have a valuable role, however, there is more to the context of language use than mere linguistic form. Therefore, it is necessary that other competences such as sociolinguistic, discourse, and strategic competence as Hymes (1972) referred to the whole pack as

communicative competence. More specifically, most of students' failure in performing tasks might have to do with their lack of knowing how to approach and implement the task.

Another important aspect of the present study was the use of the classroom's actual teacher. Most of the previous studies on this subject used either a researcher as the instructor in the study (e.g., De ridder et al., 2007; De la Fuente, 2006; Lai, Zhao, & Wang, 2011; Shintani, 2011, 2013), which would make the results of the studies biased. Other studies such as Lai et al. (2011) and Li et al. (2016) used instructors that had no prior training on TBLT, thereby making the findings of their studies questionable. However, this study used a classroom as it was with its teachers without any changes to it. Luckily, the teachers in those classes had a degree in the field of Second Language Acquisition and had already taken the TBLT courses. The use of the classroom's actual teacher in the study had two advantages, first it added to the ecological validity of the study as the classroom retained its naturalness. In fact, the class has not been alter to conduct the study in it. Second the effects of the researcher's bias on the result of the study was declined as the results of the study was of no significance to the classroom teachers.

7.2. Teachers and Classroom Research

As mentioned above, one of the encouraging points of this research study was the use of classroom teachers rather than an assigned teacher by the researchers. In fact, attempts have been made to use the same classroom teacher that the students already had in their language class for the purpose of the present study rather than the use of borrowed classrooms for the sake of research. As such, the authenticity of the research findings would be enhanced as the classroom settings were not changed to suit the hypotheses of the researchers. In the majority of the previous studies (e.g., De la Fuente, 2006; De ridder et

al., 2007; Gonzalez-Lloret & Nielson, 2015; Lai, Zhao, & Wang, 2011; Li, et al., 2016; Shintani, 2011, 2013), the role of the classroom teacher was taken for granted and those studies used either the researcher as the classroom teacher or had a different teacher who was trained for the intended methodology tested in the research. In my view, while the use of a different classroom teacher who has been especially trained to carry out the research treatment might jeopardize the authenticity of the results of the research, the use of the researcher as the classroom teacher would take the problem one step further by adding the problem of bias to the research findings. In that situation, the researcher would subconsciously be more in favor of a certain aspect of the research treatment which might lead to obtaining distorted results.

7.3. Reflection

This dissertation was originally planned to be conducted in the second language context by using the students of the English Language Institute at Florida International University. The students of this institute were taking the English language proficiency course in their preparation for starting their undergraduate studies at FIU. The researchers did the necessary coordination with the manager of the institute and the teachers to set an appointment for the collect data. However, after the data collection from two classrooms, the researchers were asked not to do the data collection during the students' actual class time. As a result, the researchers had to provide incentives for the students to be motivated to participate in the study after their regular classroom time. The researcher provided the students with the incentive of providing them with their lunch if they agreed to participate in the study. However, unfortunately the students did not show up for the data collection session except for one, which was clearly not enough for the data collection.

As a plan B, the researchers decided to collect data from the context of English as a Foreign Language from the researcher's homeland, Iran. It seemed to be a more plausible job to recruit participants from Iran as one of the researchers had friends who owned an English language institute. However, the data collection did not go as planned since some students dropped out of the study having taken the pre-tests. As such, the number of students dropped to 43 students.

Another problem with the data collection was that some students clearly did not show any interest in taking the assessments and some of them answered the items sporadically and without any commitment. As a consequence, the final number of students whose performance could be analyzed declined to 34 which was a big blow to the power of generalizability of the present study. The most important lesson learned from the present study was that human beings are not robots; they tend to choose when and how to participate in the study. Thus, the research including human subjects tends to be tentative and messy. This makes doing research a little out of control for the researchers because the research is clearly not as important to the participants of the study as to the researchers. In my view, even though this dissertation might not be methodologically perfect and flawless, it definitely had invaluable contextual authenticity and reflects the reality of working with human subjects.

7.4. Limitations

There have been a couple of limitations for this study. First, the final number of students used in this study was limited. The study began with 62 students in the first phase of the study, however, after in the posttest phase the number of students cut down to 34

and for the delayed posttest a lot of students opted not to participate in the study due to various reasons such as boredom with taking the test.

Another limitation of the study was the fact that due to the time pressure of keeping the participants of study, the tasks were altered a little so that the students would take shorter time to complete all four tests especially the test tasks. To this end, the target items in the task were underlined so that the students would be able to complete the task in a shorter time. This, per se, may have affected the taskness of the test tasks as students might have been a little overtly directed to the target feature which might per se have affected the findings of the study. That is, part of the reason that the task-based group outperformed the PPP in the accuracy of their performance might have been due to the fact the TBLT group was given a hint through the underlined items of the task. The study has been conducted using a one-hour treatment which in turn is very little for having any measurable effect especially for the acquisition of a grammatical structure to occur, especially in the case of TBLT. It seems that, due to the incidental nature of learning in TBLT, there is more treatment time needed for the students to be able to gain mastery of a certain grammatical feature.

As mentioned in the discussion part of this study, the one-hour instruction of the study as the treatment was too short to come to a robust finding about the implicit knowledge as well as the task performance competence. In fact, the explicit knowledge might well be developed through a short one-hour treatment, but this is not the case with the implicit knowledge which due to its automatic and incidental nature requires more rehearsal and practice in order to develop than a one-hour instruction. The same holds true in the case of task performance. TBLT require longer period of instruction so that it could

master the ability of learners to perform a task using a certain linguistic form. Therefore, future research should look into the possibility of using longitudinal research in order to assess the effects of TBLT and PPP on the development of implicit knowledge as well as the task performance competence.

7.6. Pedagogical Implications

The most important pedagogical implication that the present study could have for the classroom setting is that teachers should do their best to set the grounds for their students to actually learn while performing. This can well be achieved through implementing TBLT in their classroom. In essence, the ultimate goal and use of a certain language or linguistic feature is for the students to be able to use it in the real-life context; therefore, the best way that they could be able to perform and use language in their daily life is if they have already had the chance to work with a similar task in the classroom setting. In other words, classroom teachers should attempt to create an authentic and close-to-real-life situation where students could perform an authentic task. As such, the students know how to apply their knowledge of language through interaction and communication. Knowledge about the language and grammar usually promoted by PPP have shown not be as effective for in the real-life context as communicative competence promoted by TBLT. Along the same line, the teacher must use the type of assessment which is in conjunction with TBLT. In effect, students' mastery of a certain linguistic feature indicated in a discrete-point items activity would not be indicative of their successful performance of tasks in real life. As an example, a student might be able to know how to form Wh-questions in English but not be able to use those question in a restaurant ordering pizza. Therefore, TBLT is concomitant with TBLA, and only through Task Assessment teachers

can make sure that their students have gained the competence to act on their linguistic mastery to perform tasks in real-life situations.

7.7. Theoretical Implications

The present dissertation highlights the importance of Task-based Language Teaching and the advantages that it has in helping language learners master a linguistic form in the context of language use. Additionally, the study stresses the importance role of Task-Based Language Assessment in both research and instruction. In fact, Task-based Language Assessment can be a means through which students are assessed beyond their grammatical competence. Task-based Language Assessment sets the groundwork to assess what language learners are actually capable of by applying the linguistic resources they have mastered. In other words, accomplishing task assessment requires not just the mastery of the linguistic forms but showing communicative competence necessary to complete the task assessment. More specifically, as this study is a partial replication of Li et al.'s (2016) study, the findings of study reveal that the use of task assessment is necessary especially when it comes to comparing the effects of Task-based Language Teaching with any other method or approach of language teaching. Finally, in terms of the research, the present research also emphasizes the importance of using the authentic classroom teachers in conducting research as it avoids the issues of research bias and lack of authenticity.

The findings of the present study contributes to the theoretical foundations of TBLT as they explain why the use of tasks is argued to be the most effective means to facilitates noticing (e.g., Ellis, 2003; 2009; Mackey, 1999), and most critically, to *assess* learners. The theoretical premise for PPP stems from the way in which both connectionist and skill acquisition accounts of second language acquisition treat the construct of noticing

(e.g., N. C. Ellis, 2005; DeKeyser, 1998). For example, N. Ellis (2005) argues that adult language learning begins with explicit knowledge of form-meaning associations. This knowledge is developed into implicit knowledge with subsequent processing and attempts at using the language.

7.8. Suggestions for Future Research

Future research could also aim at measuring the differential effects of the PPP and TBLT instruction using focused tasks. Since focused tasks are more directed towards the linguistic form than the general tasks, it could be hypothesized that TBLT students might be better prepared to acquire a linguistic feature than when they have a general task. Since general tasks do not require the students to use the target feature, therefore, the focused task could show the TBLT students full potential by having them to perform the task only through the target feature. Additionally, the study could be conducted in the context of second language rather than foreign language as students in the context of second language acquisition have more experience of performing real-life tasks, therefore, there might be higher effects for TBLT in the context of second language. Furthermore, the future research could examine the effects of TBLT using new target structures rather than the ones that have already been partially acquired; there have been ample studies that have shown a higher effect for TBLT on structures that have already been partially acquired by the students and there is a need for more research on the effects of TBLT on new structures.

Since the TBLT treatment in this study followed Willis's model, future research could use the other model of TBLT implementation, being the Ellis model. The main difference between the Ellis model and Willis's model is the phase in which the

grammatical review or instruction is provided. In fact, in the Ellis model, the grammatical review or instruction could be offered at any of the three phases of pre-, during-, and post-task. However, in Willis's model, the grammatical review could only be provided at the language focus phase. Therefore, conducting research using either of these two models could yield different results.

It would be interesting if future research could be conducted using similar research in the context of second language. The present study in its earlier phases set to assess the effects of TBLT and PPP on students' performance in the context of second language; however, due to the limitations of sources and lack of participants in the US, the study was conducted in a foreign language context. The reason for this suggestion is that usually TBLT tends to yield even better results with learners in the second language context as they have experience with the use of language in an authentic context. Therefore, there might be some advantage on the part of TBLT students if the present study was done in the context of second language, and maybe better results for the TBLT students.

The results of the study might have been affected by the cultural aspect of the research setting. The dominant use of the PPP approach in the Iranian contexts over decades might have affected how comfortable the students were with the PPP approach. Therefore, if the same research is done in a different cultural context in a different country, the results might well be different. Future research could replicate the same study in a context where the cultural acceptance of these two approaches is different.

REFERENCES

- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Bachman, L. F. (2002). Some reflections on task-based language performance assessment. *Language Testing*, 19, 453-476.
- Bachman, L. F. (2007). What is the construct? The dialectic of abilities and contexts in defining constructs in language assessment. In J. Fox, M. Wesche, D. Bayliss, L. Cheng, C. Turner, & C. Doe (Eds.), *Language testing reconsidered* (pp. 41–71). Ottawa, Canada: University of Ottawa Press.
- Baker, D. (1989). *Language testing: A critical survey and practical guide*. London: Edward Arnold.
- Baralt, M., & Morcillo Gomez, J. (2017). Task-based language teaching online: A guide for teachers. *Language Learning & Technology*, 21(3), 28-43.
- Baralt, M., Gilabert, R., & Robinson, P. (2014). *Task sequencing and instructed second language learning*. London: Bloomsbury.
- Brindley, G. (1994). Outcomes-based assessment and reporting in language learning programmes: A review of the issues. *Language Testing*, 15, 45-85.
- Brindley, G. (2013). TBLA. In Chapelle, C. (ed.), *The encyclopedia of applied linguistics*. Oxford: Wiley Blackwell.
- Brown, J. D. (2004). Performance assessment: Existing literature and directions for research. *Second Language Studies*, 22(2), 91-139.
- Carless, D. (2004). Issues in teachers' reinterpretation of a task-based innovation in primary schools. *TESOL Quarterly*, 38(4), 639-662.
- DeKeyser, R. (1998). Beyond focus on form: Cognitive perspectives on learning and practicing second language grammar. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 42–63). Cambridge, UK: Cambridge University Press.
- DeKeyser, R. (2007). Skill acquisition theory. In J. Williams & B. VanPatten (Eds.), *Theories in second language acquisition: An introduction* (pp. 97–113). Mahwah, NJ: Erlbaum.
- De La Fuente, M., J. (2006). Classroom L2 vocabulary acquisition: Investigating the role of pedagogical tasks and form-focused instruction. *Language Teaching Research*, 10, 263-95.
- De Ridder, I., Vangehuchten, L., & Gómez, M. S. (2007). Enhancing automaticity through task-based language learning. *Applied Linguistics*, 28, 309-15.

- Ellis, N. C. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in second language acquisition*, 27(2), 305-352.
- Ellis, R. (1991). Communicative competence and the Japanese learner. *JALT Journal*, 13, 103-29.
- Ellis, R. (2002). The place of grammar instruction in the second/foreign language curriculum. In Hinkel, E., & Fotos, S. (eds.), *New perspectives on grammar teaching in second language classrooms* (pp. 17–34). Mahwah, NJ: Lawrence Erlbaum.
- Ellis, R. (2003). *Task-based language learning and teaching*. Oxford: Oxford University Press.
- Ellis, R. (2004). The definition and measurement of L2 explicit knowledge. *Language Learning*, 54, 227–75.
- Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language: A psychometric study. *Studies in Second Language Acquisition*, 27(2), 141–172.
- Ellis, R. (2009). Task-based language teaching: Sorting out the misunderstandings. *International Journal of Applied Linguistics*, 19, 221-46.
- González-Lloret, M., & Nielson, K. B. (2015). Evaluating TBLT: The case of a task-based Spanish program. *Language Teaching Research*, 19, 525-49.
- Hollander, M., & Wolfe, D. A. (1999). *Nonparametric Statistical Methods*. New York City, NY: Wiley.
- Hymes, D. H. (1972). On Communicative Competence. In Pride, J. B., & Holmes, J. (Eds.), *Sociolinguistics*, 269-293. Baltimore, USA: Penguin Education, Penguin Books Ltd.
- Krashen, S. (1982). *Principles and practice in second language acquisition*. New York: Pergamon.
- Lai, C., Zhao, Y., & Wang, J. (2011). Task-based language teaching in Online Ab Initio foreign language classrooms. *The Modern Language Journal*, 95, 81-103.
- Li, S., Ellis, R., & Zhu, Y. (2016). Task-based versus task-supported language instruction: An experimental study. *Annual Review of Applied Linguistics*, 36, 205–29.
- Long, M. H. (1983). Native speaker/non-native speaker conversation and the negotiation of comprehensible input. *Applied Linguistics*, 4, 126–141.
- Long, M. (1985). A role for instruction in second language acquisition: task-based language teaching. In K. Hyltenstam & M. Pienemann (Eds.), *Modelling and assessing second language acquisition*. Clevedon: Multilingual Matters.
- Long, M. H. (1989). Task, group, and task-group interactions. *University of Hawai'i Working Papers in ESL*, 8, 1–25.
- Long, M. H. (1991). Focus on form: A design feature in language teaching methodology. In K. de Bot, R. Ginsberg, & C. Kramersch (Eds.), *Foreign language research in cross-cultural perspective* (pp. 39–52). Amsterdam: Benjamins.

- Long, M. H. (2009). Methodological principles in language teaching. In Long, M.H., & Doughty, C.J. (eds.), *Handbook of language teaching* (pp. 373–394). Oxford: Blackwell.
- Long, M. (2015). *Second language acquisition and task-based language teaching*. Malden, MA: Wiley Blackwell.
- Long, M. H., & Norris, J. M. (2000). Task-based teaching and assessment. In M. Byram (Ed.), *Encyclopedia of language teaching* (pp. 597-603). London: Routledge.
- Long, M. H., & Robinson, P. (1998). Focus on form: Theory, research, and practice. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 15–41). New York: Cambridge University Press.
- Mackey, A. (1999). Input, interaction, and second language development: An empirical study of question formation in ESL. *Studies in second language acquisition*, 21(4), 557-587.
- Mislevy, R., Steinberg, L., & Almond, R. (2002). Design and analysis in task-based language assessment. *Language Testing*, 19(4), 477–496.
- Nassaji, H., & Fotos, S. (2011). *Teaching grammar in second language classrooms: Integrating form-focused instruction in communicative context*. New York: Routledge.
- Norris, J. M. (2016). Current Uses for Task-Based Language Assessment. *Annual Review of Applied Linguistics*, 36, 230-244.
- Nunan, D. (2004). *Task-based language teaching*. Cambridge: Cambridge University Press.
- Norris, J., Brown, J., Hudson, T., & Yoshioka, J. (1998). *Designing second language performance assessments*. Honolulu, HI: University of Hawai‘i Press.
- Norris, J. M. (2016). Current uses for task-based language assessment. *Annual Review of Applied Linguistics*, 36, 230-244.
- Richards, J. C., & Rodgers, T. S. (2014). *Approaches and methods in language teaching*. Cambridge: Cambridge university press.
- Robinson, P. (1996). Task-based testing, performance-referencing and program development. *The University of Queensland Working Papers in Language and Linguistics*, 1, 95-117.
- Robinson, P. (2011). Task-based language learning: A review of issues. *Language Learning*, 61, Suppl. 1, 1–36.
- Robinson, P., & Ross, S. (1996). The development of TBLA in English for academic purposes programs. *Applied Linguistics*, 17, 4, 455–476.
- Samuda, V., & Bygate, M. (2008). *Tasks in second language learning*. London: Palgrave Macmillan.

- Seedhouse, P. (1999). Task-based interaction? *ELT Journal*, 53, 149–156.
- Seedhouse, P. (2005). “Task” as research construct. *Language Learning*, 55(3), 533-570.
- Sheen, R. (1994). A critical analysis of the advocacy of the task-based syllabus. *TESOL Quarterly*, 28(1), 127-151.
- Sheen, R. (2003). Focus on form—a myth in the making? *ELT Journal*. 57(3), 225-233.
- Skehan, P. (1996). Second language acquisition research and task-based instruction. In J. Willis & D. Willis (Eds.), *Challenge and change in language teaching* (pp. 17-30). Oxford: Macmillan Heinemann.
- Skehan, P. (1998). Task-based instruction. *Annual Review of Applied Linguistics*, 18, 268–86.
- Shintani, N. (2011). A comparative study of the effects of input-based and production-based instruction on vocabulary acquisition by young EFL learners. *Language Teaching Research*, 15, 137-58.
- Shintani, N. (2013). The effect of focus on form and focus on forms instruction on the acquisition of productive knowledge of L2 vocabulary by young beginning-level learners. *TESOL quarterly*, 47, 36-62.
- Swan, M. (2005). Legislation by hypothesis: The case of task-based instruction. *Applied Linguistics*, 26(3), 376-401.
- Van den Branden, K. (Ed.). (2006). *Task-based language education: from theory to practice*. Cambridge: Cambridge University Press.
- Van den Branden, K. (2009). Mediating between predetermined order and chaos: the role of the teacher in task-based language education. *International Journal of Applied Linguistics*, 19, 264-285.
- Widdowson, H. (2003). *Defining issues in English language teaching*. Oxford: Oxford University Press.
- Willis, J. (1996). *A framework for task-based learning*. London, UK: Longman.
- Willis, D., & Willis, J. (2007). *Doing task-based teaching*. Oxford: Oxford University Press.
- Williams, M., & Burden, R. L. (2015). *Psychology for language teachers: a social constructivist approach*. Cambridge: Cambridge University.

APPENDICES

**Appendix A:
The Grammaticality Judgement Test**

Please read the sentences below. Mark 'C' for the sentences that are correct 'I' for the sentences that are incorrect. For any sentence that you think is incorrect, cross out the error and write the correct version on the line below. Please write down the correct form of those items that are incorrect on the lines below them.

Example:

Yesterday I goed to the store. I
went

1. The show was repeat twice last month. _____

2. He decided to learn the Spanish Language. _____

3. The keys find on the back seat of the taxi. _____

4. The best player in the team badly injured in the game. _____

5. The thief was arrest the following day. _____

6. He bought an expensive car for his son. _____

7. 70,000 people killed in the Wenchuan earthquake in 2008. _____

8. Mary was give a mobile phone for her birthday. _____

9. The boy take to the headmaster for questioning. _____

10. Some new flowers were plant in the garden. _____

11. He promised to wash the dishes. _____

12. The gold buried under a big tree in 1900. _____

13. That young woman raised in a large rich family. _____

14. They have sent me a card for Christmas. _____

15. Yesterday food and clothes bring to help the people. _____

16. This morning Helen was knock down in the street. _____

17. John's knee seriously hurt in a cycling event. _____

18. She works at a hospital in a small town. _____

19. Many questions were discuss at our last meeting. _____

20. The school building was paint red in 1970. _____

21. His leg was break in a practice game before the sports meeting. _____

22. They will travel to California next month. _____

23. The students were tell to listen carefully to the teacher. _____

24. The new student's name added to the class list. _____

25. Mary's foot cut on her way home from school. _____

26. She doesn't know the Spanish language. _____

27. Yesterday books collect to help the poor children. _____

28. Last week the old bike repair in a bicycle shop. _____

29. The lantern put on a table in the corner. _____

30. It snows a lot in winter in their city. _____

31. After the accident, the victims treat in a local hospital. _____

32. These songs recorded over two years ago. _____

33. The tall building destroy during the flood. _____

34. My doctor told me to stop smoking. _____

35. This morning a special key use to open the door. _____

36. Last month all the parents invited to the meeting. _____

37. She exercises four times a week to lose weight. _____

38. The beautiful house damage in a snowstorm. _____

39. Tony badly hit in a fight with a friend. _____

40. The students were allow to stay in the library. _____

Appendix B:

The Elicited Imitation Test

Listen to the recording of 35 item. After hearing each sentence, say if it is true of you or not and then state the sentence in correct English.

1. One of my friends was killed last week. _____
2. My knee injure on my way to school today. _____
3. My grandfather treated at a hospital last week. _____
4. My lovely bicycle was damage last week. _____
5. **The quality of food in this restaurant is high.** _____
6. A student from my school was arrested yesterday. _____
7. A window was break in my house today. _____
8. My finger was very badly cut this morning. _____
9. I was tell to hand in my homework yesterday. _____
10. My father was hit in a car accident last year. _____
11. I was take to a dentist a few days ago. _____
12. **I will do the assignments tomorrow.** _____
13. My father raise in a poor family. _____
14. I was knocked down in the street yesterday. _____
15. A bridge near my house destroyed last year. _____
16. My grandmother was bury in her village last year. _____
17. I was allowed to watch TV last weekend. _____
18. My friend was badly hurt in a fight yesterday. _____
19. My father was brought home drunk yesterday. _____
20. My grandfather was find dead in the street yesterday. _____

21. My painting was put up on the school wall last month. _____
22. **I have been to many countries in Europe.** _____
23. I was give a nice present on my birthday. _____
24. My English presentation was recorded last time. _____
25. I was add to the school soccer team this year. _____
26. My parents were invite to a dinner during the weekend. _____
27. **He decided to save his money for the trip.** _____
28. Garbage was collected from our dormitory room last Monday. _____
29. Our family car used for shopping last weekend. _____
30. My watch repaired in a local shop yesterday. _____
31. My favorite program was repeated on TV this week. _____
32. **I bought my mom a gift on Mother's Day.** _____
33. My house paint white last year. _____
34. Some trees were planted at my school last year. _____
35. Many interesting things were discussed in my English class today.

Appendix C:

Task Assessment

Part 1. Read the story below about a robbery that took place in Miami last month.

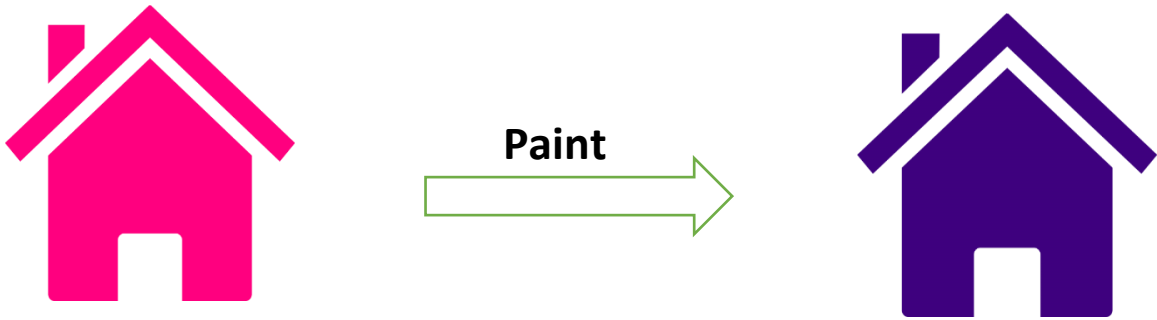
Yesterday morning some people called the police and gave the police the information that two armed men robbed a bank. The two men stole about 2 million dollars from the bank. The robbers hurt some customers and shot two bank clerks. They also broke the windows of the bank in the meantime. The cameras in the bank filmed the robbers. When the police arrived, they asked the people around the bank to leave the scene. The police hit one of the two robbers. The robbers set fire to a car when they were escaping. The police shot one robber in the shoulder but the other robber helped him escape. The police did not find the robber yet. The robber left some clues at the crime scene. The police interviewed the bank clerks and the customers for more information. The television showed videos of the robbery and people spread the news of the robbery in the town.

Part 2. Now, rewrite the story using the passive voice. Then, add your own creative ending to your story.

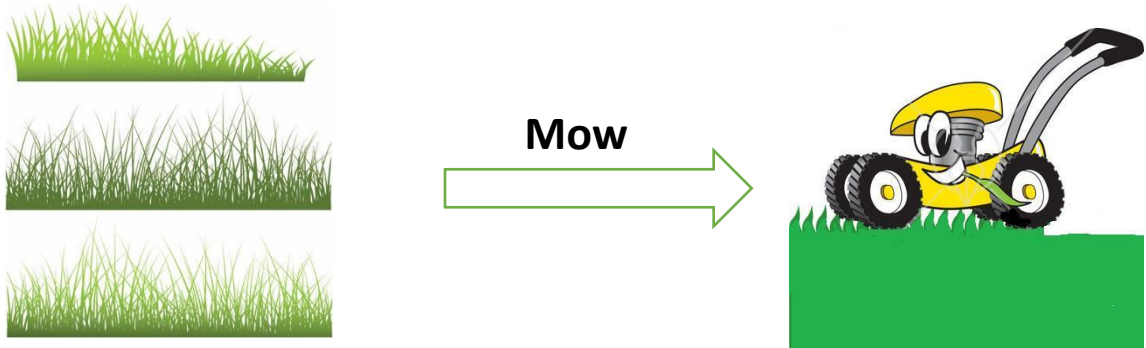
Appendix D: The TBLT Treatment Material

Treatment Task 1

Rebecca's brother has moved away, and a lot has changed since then. Rebecca is going to write her brother a letter explaining what changes she has made since he moved. On the sheet of paper, please help Rebecca write her letter and explain the changes. You will write eight changes total. Use the pictures below to help you.



1. The house



2. The grass/lawn



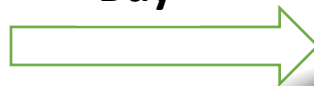
Plant



3. The tree



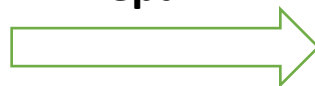
Buy



4. A new car



Repair



5. The washing machine



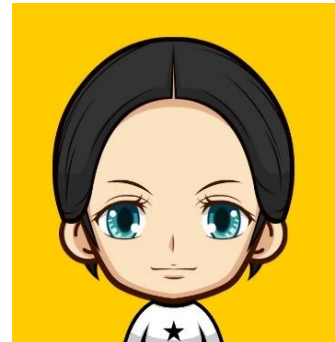
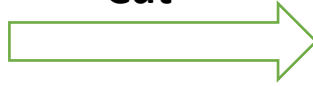
Change/Fix



6. The light bulb



Cut



7. My hair



Arrange



8. The books



Clean →



9. Your room

Treatment Task 2

Part 1. Read Cindy's mom's note on the table about what her family was supposed to do while she would be away for some days.

"Dear Cindy,

My flight is today at 5 o'clock and as you know I'm heading for the airport now. I want you to do some stuff while I'm away. First of all, water the flowers in the sitting room.

Take the garbage out on Monday night. Don't forget to give the cats their food. Give the puppy her medicine every morning or she'd get sick again. Her medicine is in the

sitting room next to the TV. Take the car to the garage or be stuck in the middle of the day. I put some money in the drawer, it's my payment money. Go to the bank and pay

the money. I have also made you lunch, you need to heat it up; it's your favorite dish, pasta. Please clean the house, your dad is allergic to cats and dogs. He'd get sick if you

don't clean the room of their fur. One last thing, do the dishes and don't leave any in the sink. Take care of little Sarah and help her with her test. I love you all. Take good

care of yourself."

Kisses,
Mom

Part 2. Now, Cindy needs to respond to her mom by leaving her a text. In her text, Cindy wants to reassure her mother that she has done everything that her mother asked her to do. Complete her text message below.

Hi mom! Done! - I did all that you asked me to do. The flowers were _____, the garbage _____, and _____.

Also, the medicine _____, the car _____, and the money

_____. Besides, the lunch _____, the house

_____, and the dishes _____. Also, Sarah

_____. See you soon – I love you!

Treatment Task 3

Part 1. Read the passage below about John, his father, and their family's timeless tradition of making bread. Try your best to remember the story, because when you are done, you will have to put the story's sequence of events in the same order that they appeared in the story.

Here is John and his dad's conversation on breakfast table.

John: Dad, I have always wondered how this bread is produced.

Dad: well, if you mean this particular bread. This is actually made of wheat.

John: how is it made of wheat?

Dad: well, first the farmer made the land ready for **planting the wheat**. This was done by **watering the land** so that the soil was ready for planting. When the soil of the land was soft and ready after watering it, the farmer used a tractor to dig the land and **make it ready** for planting the wheat seeds in the soil. After that, the wheat seeds received sunlight and grew. When the seeds grew, they turned into wheat plants. Then, the farmer used a machine **to cut** the wheat plants and after that the machine **separated** the wheat from the chaff. Then, the separated wheats **were taken** to the factory. In the factory, a machine **crushed** the wheat into a powder. Then the powder **was sent** to the market. The bakeries **bought** the powder and **added** water and some other substance to the wheat powder to turn it to the dough. The bakeries cut the dough into pieces and **kneaded** it to make it ready for baking. The last phase was when **the dough was baked** and the bread is ready.

John: wow, such a long process.

Dad: yes, a simple thing such as bread takes so long to be produced.

Part 2. Look at the sentences below. Rewrite the process of making a bread from the wheat by putting these sentences in the right order.

- _____ The land **was watered** by the farmer or the rain.
- _____ The dough **was kneaded** by the machine.
- _____ Water and other substances **were added** to wheat powder.
- _____ The wheat seeds **were planted** by the farmer.
- _____ The dough was baked.
- _____ The wheat powder was sent to the market.
- _____ The wheat was crushed by the machine to be a powder.
- _____ The dough was shaped and cut into pieces.
- _____ The wheat plants **were cut** by the machine.
- _____ The wheat seeds **grew and turned** into wheat plant over time.
- _____ The dough was baked and tuned into bread.
- _____ The land soil **was made** ready by the tractor.
- _____ The wheat seeds **were put** to the sunlight.
- _____ After being crushed, the wheat **was turned** to the wheat flour.
- _____ The wheat was taken to the factory.
- _____ The wheat flour **was turned** to a dough.
- _____ The wheat were separated from the chaff.
- _____ The wheat powder was sent by the farmer.

Appendix E: The PPP Treatment

Presentation (20 minutes)

The students will be given some explanations regarding the structure of the past passive through ample use of examples. First the past form of the verb to be is explained.

The present form of to be	The past form of to be
am/is	was
are	were

Examples:

I **am** at school now.

I **was** at school yesterday.

They **are** in the gym today.

They **were** in the gym yesterday.

Then, the students are explained the role of the past participle (PP.) of a verb. There are two different types of the past participle form of a verb: one for the regular verbs, the other for irregular verbs. The past participles of the regular verbs is formed by adding an 'ed' to the end of the verb, while the past participle of the irregular verbs does not follow any rules and must be learned separately.

Example: (Regular verbs)

The present form of the verb	The past form of the verb	The past participles form of the verb
Play	Played	Played
Help	Helped	Helped
Kill	Killed	Killed
Ask	Asked	Asked
Like	Liked	Liked
Park	Parked	Parked

Example: (Irregular verbs)

The present form of the verb	The past form of the verb	The past participles form of the verb
Go	Went	Gone
Do	Did	Done
See	Saw	Seen
Take	Took	Taken
Give	Gave	Given
Run	Ran	Run
Come	Came	Come
Bring	Brought	Brought
Tell	Told	Told
Make	Made	Made

After explaining the different form of past participle, the full structure of the past passive form is introduced.

Subject + simple past of TO BE + past participle

Subject + simple past of TO BE + past participle + by + agent

Examples:

The cat **killed** the mouse.

The mouse **was killed** by the cat.

My friend **brought** the cake.

The cake **was brought** by my friend.

Practice: (30 minutes)

In this phase, the students will be given some time to complete drills about past passive structure.

1. Complete the sentences with the correct form of the verbs in the parentheses. Use the past passive tense.

1. The cars(make) in USA.
2. Where(building/make)?
3. The song(sing) by Britney Spears.
4. The elephant(keep) in the zoo.
5. When(food/cook)?
6. How many (people/shoot)?
7. When(car/buy)?
8. Her nails(paint/red).
9. The match (cancel) yesterday.
10. The ruler(break) by John.

2. Choose the correct form of the verbs in brackets.

Fiat 0 was started (started/was started) by a group of Italian businessmen in 1899. In 1903, Fiat, 1 _____ (produced/was produced) 132 cars. Some of these cars 2 _____ (exported/were exported) by the company to the United States and Britain. In 1920, Fiat 3 _____ (started/was started) making cars at a new factory at Lingotto, near Turin. There was a track on the roof where the cars 4 _____ (tested/were tested) by technicians. In 1936, Fiat launched the Fiat 500. This car 5 _____ (called/was called) the Topolino – the Italian name for Mickey Mouse. The company grew, and in 1963 Fiat 6 _____ (exported/was exported) more than 300,000 vehicles. Today, Fiat is based in Turin, and its cars 7 _____ (sold/are sold) all over the world.

3. Change the following sentences into passive sentences using the words in brackets.

a. We sold tickets for all shows at the Box Office. (Tickets for all shows/sell/at the Box Office)

b. Thomas Edison invented the electric light bulb. (The electric light bulb/invent/by Thomas Edison)

c. Someone painted the office last week. (The office/paint/last week)

d. Several people saw the accident. (The accident/see/by several people)

e. Where did they make these video recorders? (Where/these video recorders/make)

Production: (30 minutes)

In this phase, the teacher makes use of the following controlled drills in order to practice the chunks of language will students

1. Repetition:

Example:

Teacher: the injured was taken.

Students: the injured was taken.

Teacher: the injured was taken to the hospital.

Students: the injured was taken to the hospital.

Teacher: the injured was taken to the hospital by the police.

Students: the injured was taken to the hospital by the police.

1. The food was made by my mother.
2. The car was fixed by the repairman.
3. The house was built by the construction company.
4. The money was stolen by the thieves.
5. The trees was broken down by the storm.
6. The suitcase was loaded onto the car by my brother.
7. The cake was ruined by the guests.
8. The glass was broken by children.
9. The soldier was killed by the enemy.
10. The car was washed by that man.

2. Substitution

Example.

Teacher: the houses was destroyed by the wind. (they)

Students: they were destroyed by the wind.

1. The fire was out by the firefighter. (it)
2. The cats were fed by Sarah. (they)
3. The car was painted by him. (it)
4. The antlers were killed by the lion. (they)
5. The glass was shattered by my sister. (it)
1. The lawn was mowed by that man. (him)
2. The cake was baked by my mother. (her)
3. The milk was drunk by Susie. (her)

4. The cigarette was smoked by John. (him).
5. The wood was burned by the farmers. (them)

3. Restoration.

1. Students/help/teacher
2. Toy/buy/her father
3. Man/cure/doctor
4. Building/paint/painter
5. Bomb/explode/terrorist
6. Power/invent/Edison
7. Honey/produce/bees
8. Kitchen/clean/my sister
9. Frog/attack/snake
10. President/interviewed/journalist

4. Backwards Build-Up

Example:

Teacher: by the butcher

Students: by the butcher

Teacher: was cut by the butcher

Students: was cut by the butcher

Teacher: the meat was cut by the butcher.

Students: they meat was cut by the butcher.

1. The tea was prepared by my aunt.
2. The clothes were washed by the washing machine.
3. The stone was thrown at the boy by his friend.
4. The girl was adopted by the parents.
5. The book was bought by my father.
6. The air was polluted by the truck.
7. My hair was cut by the barber.
8. The new clothes were worn by the students.
9. The new lesson was taught by the teacher.

Appendix F: Control Group Treatment

6



Peter Judd joined the army when he was eighteen, and for several months he was taught how to be a good soldier. He did quite well in everything except shooting. One day he and his friends were practising their shooting, and all of them were doing quite well except Peter. After he had shot at the target nine times and had not hit it once, the officer who was trying to teach the young soldiers to shoot said, 'You're quite hopeless, Peter! Don't waste your last bullet too! Go behind that wall and shoot yourself with it!'

Peter felt ashamed. He went behind the wall, and a few seconds later the officer and the other young soldiers heard the sound of a shot.

'Heavens!' the officer said. 'Has that silly man really shot himself?'

He ran behind the wall anxiously, but Peter was all right. 'I'm sorry, sir,' he said, 'but I missed again.'

A Which of these sentences are true (T) and which are false (F)? Write T or F in the boxes.

1. Peter was good at everything.
2. Peter was not good at shooting.

Outside the 1 500 headwords: shot (n.), target

3. A lot of the other soldiers were bad at shooting too.
4. One of Peter's nine bullets hit the target.
5. The officer was not pleased with him.
6. The officer thought that Peter had shot himself.

B Write these sentences. Choose the correct word in each:



1. Peter is looking at an enemy and trying to shoot him. himself.



2. Now Peter is trying to shoot him. himself.



3. Peter's girl-friend is sitting in front of her sister and making her herself up.



4. Peter's girl-friend is sitting in front of her sister and making her herself up.



5. The white donkey is in a field with a black donkey, and it has hurt it. itself.



6. The white donkey is in a field with a black one, and it has hurt it. itself.



7. The boys have brought their small sisters to a shop and bought them themselves some sweets.



8. The boys have brought their small sisters to a shop and bought them themselves some sweets.



Mr Richards worked in a small seaside town, and he and his wife had a comfortable house near the sea. During the winter they were quite happy there, but every summer a lot of their relatives used to want to come and stay with them, because it was a nice place for a holiday, and it was much cheaper than staying in a hotel.

Finally one June Mr Richards complained to an intelligent friend of his who lived in the same place. 'One of my wife's cousins intends to bring her husband and children and spend ten days with us next month again. How do you prevent all *your* relatives coming to live with you in the summer?'

'Oh,' the friend answered, 'that isn't difficult. I just borrow money from all the rich ones, and lend it to all the poor ones. After that, none of them come again.'

A Which of these sentences are true (T) and which are false (F)? Write T or F in the boxes.

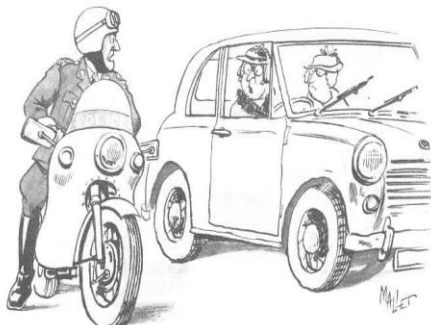
1. Mr and Mrs Richards did not have many relatives.
2. Mr and Mrs Richards often went to stay with their relatives in the summer.
3. Mrs Richards's cousin decided to visit them.
4. Mr Richards was not at all happy about this.
5. A friend of his told him how to stop visitors.
6. He borrowed money from his poor relatives, and lent money to his rich ones.

B Answer these questions:

1. Where did Mr and Mrs Richards live?
2. What was their house like?
3. Why were they happier in winter than in summer?
4. Why did their relatives want to visit them?
5. Who did Mr Richards ask for advice then?
6. What did he say to him?
7. And what did his friend answer?
8. Why did the friend's relatives not want to visit him again?

C Put one word in each empty place. You will find all the words in the story on page 16.

George liked the sea, so he lived in a house at the His parents and a lot of his other . . . lived near him. George's chairs were very hard, so they were not very . . . to sit on. . . . one day his mother . . . about this, saying, 'I wish you had some softer chairs, George.' But his father laughed and answered, 'George is an . . . man: he has hard chairs because he . . . to . . . people . . . too long when they come to visit him! When he wants a soft chair, he can . . . it from our house.'



Helen lived with her sister Mary. Both of them were about seventy-five years old, and neither of them had ever married. They had a small, old car, and when they wanted to go somewhere, which they did very rarely, Mary always drove, because her eyes were better.

One weekend they drove to a large town to look at some things which they had read about in the newspaper. Neither of them had been to that town before.

They were driving along in a lot of traffic when they turned right into a street which cars were not allowed to go into. There was a policeman there, and he blew his whistle, but Mary did not stop, so he got on to his motor-cycle and followed them.

After he had ordered them to stop, he said, 'Didn't you hear me blow my whistle?'

'Yes, we did,' admitted Mary politely, 'but Mummy told us never to stop when men whistle at us.'

A Which of these sentences are true (T) and which are false (F)? Write T or F in the boxes.

1. Helen and Mary were sisters.
2. They were quite young.
3. They were both married.
4. Mary drove their car.
5. A policeman tried to stop her one day.
6. She did not stop, because she did not hear him blow his whistle.

B Answer these questions:

1. Where did Mary and Helen drive one day?
2. What did they do in the town?
3. What did the policeman do?
4. And what did Mary do?
5. What did the policeman do then?
6. What did he say to Mary when he stopped her?
7. And what did Mary answer?
8. Had her mother really meant that she should not stop when a policeman blew his whistle?

C Find words in the story on page 20 which mean about the same as:

1. at any time
2. big
3. cars, buses, vans, etc.
4. confessed
5. got husbands
6. Mother
7. not rudely
8. Saturday or Sunday
9. seldom
10. told
11. went after



In England nobody under the age of eighteen is allowed to drink in a public bar.

Mr Thompson used to go to a bar near his house quite often, but he never took his son, Tom, because he was too young. Then when Tom had his eighteenth birthday, Mr Thompson took him to his usual bar for the first time. They drank for half an hour, and then Mr Thompson said to his son, 'Now, Tom, I want to teach you a useful lesson. You must always be careful not to drink too much. And how do you know when you've had enough? Well, I'll tell you. Do you see those two lights at the end of the bar? When they seem to have become four, you've had enough and should go home.'

'But, Dad,' said Tom, 'I can only see one light at the end of the bar.'

A Which of these sentences are true (T) and which are false (F)? Write T or F in the boxes.

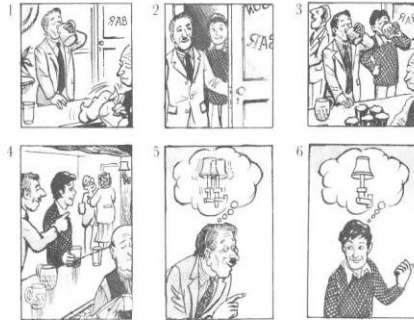
1. When Tom was under eighteen, his father took him to a public bar.
2. It was the first time that he had taken him to his usual bar.
3. There was one light at the end of the bar.

4. Mr Thompson wanted to teach Tom not to drink too much.
5. Mr Thompson thought he saw four lights.
6. Tom only saw two.

B Answer these questions:

1. Who can drink in public bars in England?
2. Why did Mr Thompson not take Tom to his usual bar for a long time?
3. When *did* he take him?
4. What did they do there?
5. What did Mr Thompson say then?
6. And what did Tom answer?
7. Who had had enough to drink, Tom or his father?
8. How did Tom know this?

C Put the correct sentences under the correct pictures:



1. But Tom could see only one.
2. He thought he could see two lights.
3. Mr Thompson used to go to a bar alone.
4. Then Mr Thompson pointed to the light at the end of the bar.
5. Then, when Tom was eighteen, he took him to the bar too.
6. They drank beer.



Jim was intelligent, but he hated hard work. He said, 'You work hard, and earn a lot of money, and then the government takes most of it. I want easy work that gives me lots of money and that the government doesn't know about.'

So he became a thief—but *he* did not do the stealing: he got others to do it. They were much less intelligent than he was, so he arranged everything and told them what to do.

One day they were looking for rich families to rob, and Jim sent one of them to a large beautiful house just outside the town.

It was evening, and when the man looked through one of the windows, he saw a young man and a girl playing a duet on a piano.

When he went back to Jim, he said, 'That family can't have much money. Two people were playing on the same piano there.'

Outside the 1 500 headwords: duet, rob

52

A Which of these sentences are true (T) and which are false (F)? Write T or F in the boxes.

1. Jim was clever, but lazy.
2. He didn't like paying taxes.
3. He always used clever people to steal for him.
4. They tried to find rich families.
5. One of his men played a duet on a piano with a girl.
6. The girl's family was poor.

<input type="checkbox"/>
<input type="checkbox"/>
<input type="checkbox"/>
<input type="checkbox"/>
<input type="checkbox"/>
<input type="checkbox"/>

B Answer these questions:

1. Why did Jim become a thief?
2. How did he steal things?
3. What was *his* job in this?
4. Why was he good at it?
5. What did he send one of his men to do one evening?
6. What did the man see?
7. What did he say to Jim?
8. What mistake did he make?

C Opposites: Find words in the story on page 52 which mean about the opposite of:

1. difficult
2. lazy
3. little
4. more
5. old
6. poor
7. small
8. spend
9. stupid
10. ugly

53

VITA

MAJEED NOROOZI

Born, Mazandaran, Iran

2003-2007 B.A., English Language Translation
Payamenour University
Mazandaran, Iran

2007-2010 M.A., Teaching English
University of Mazandaran
Mazandaran, Iran

2013-2014 Fulbright Language Scholar
Miami, Florida

2014-2018 Doctoral Candidate
Florida International University
Miami, Florida

Teaching Assistant
Florida International University
Miami, Florida

PUBLICATIONS AND PRESENTATIONS

Marzban, A., & Noroozi, M. (2012). Strategic planning focus and time in narrative writing. *Procedia - Social and Behavioral Sciences*, 46, 4953–4956.

Noroozi, M. (2010, October). *Strategic planning focus and time in narrative writing*. Paper presented at the 8th International TELLSI Conference, CALL for change in our Language Teaching, TELLSI, Tehran, Iran.