Improving EFL Learners' Oral Production through Reasoning-gap Tasks Enhanced with Critical Thinking Standards: Developing and Implementing a Critical TBLT Model, Pre-Task Plan, and Speaking Rubric

Zeynep Yaprak[1*], Fatma Kaya

[1]*The Department of Foreign Language Education, Dicle University 21280 Sur/Diyarbakır, Turkey*
[2]*The Department of Foreign Language Education, Dicle University 21280 Sur/Diyarbakır, Turkey*

Corresponding Author: Zeynep Yaprak, E-mail: zeynep.yaprak@dicle.edu.tr

ARTICLE INFO

ABSTRACT

This study aims to design and implement a speaking task model following the principles of Task-Based Language Teaching (TBLT) and Critical Thinking (CT). The study mainly investigated to what extent the use of critical thinking standards had an impact on the students' oral performance with respect to the linguistic trilogy of fluency, accuracy, and complexity. Experimental research design was conducted to examine the differences in research foci between the non-treatment and treatment groups. Data were drawn from 16 students of the English Language Teaching department and collected through the designed task rubric, speech samples, and semi-structured interview protocols. The findings indicated that adherence to relevant intellectual standards in reasoning-gap tasks had a positive effect on the oral performance of the speakers of the target language. Additionally, concerning two linguistic dimensions of accuracy (p=.00) and complexity (p=.00), MANOVA results showed a statistically significant difference between two main research groups. Based on the semi-structured interview findings, all participants in the treatment group reported positive views regarding their learning experiences and oral performance supported with critically enhanced reasoning-gap tasks. This study proposes language teachers a synthesized TBLT model enhanced with critical thinking standards. With this task model, suggested pre-task plan, and speaking rubric, they can boost their students' intellectual agility and make them more intellectually and linguistically active in their oral performance.

INTRODUCTION

With the emergence of trends in communicative language teaching that aim to design language classrooms to be more interaction-oriented and more learner-centered, task-based language teaching has become an indispensable notion in second language teaching. However, some erroneous beliefs, such as purely aiming for "immediate communication" without taking interlanguage development into account and emphasizing "lexical modes of communication", have caused misunderstandings and non-productive applications of task-based approaches in the second language teaching and learning process (Bruton, 2005; Shehadeh, 2005; Skehan, 1996). On the other hand, for more productive outcomes and more active learner participation, which is more conducive to second language acquisition, language needs to be seen as "a meaning system" as well (Shehadeh, 2005). The critical warnings against pure structuralism and the tendency to overlook the value of meaningful use of language have given rise to significant and innovative practices in second language education. In the quest for more authentic, meaningful language production, and learner engagement, TBLT provides opportunities for learners to practice the target language through real-world tasks (Ellis, 2003; Foster and Skehan, 1996; Nunan, 1989; Prabhu, 1987). Activities in language teaching need to meet some criteria to be specified as a "task". According to Ellis (2009), these criteria are:

1. Focus-on meaning should be the main concern.
2. There needs to be "some kind of gap" through which learners should be encouraged and pushed to transfer information, express their opinions, and deduce meaning from evidence.
3. It is essential to ensure that learners apply their own linguistic and non-linguistic resources.
4. There is a specific outcome other than the display of language.

Considering these standards, it can be indicated that TBLT, most importantly, promotes purposeful communication through developing learners' both linguistic and cognitive competence. In addition to their linguistic function, by

placing learners in real-world situations, language tasks play a key role in promoting learners' ability to think and reason. However, making people think is not easy, or enough, as "[e]veryone thinks; it is our nature to do so. But much of our thinking, left to itself, is biased, distorted, partial, uninformed or down-right prejudiced" (Paul and Elder, 2009). For that reason, raising the students' awareness of critical thinking and explicit instruction in critical thinking skills could be effective ways to support the learners' intellectual engagement in language learning. Within this interrelationship between language and intellectual development, for more accurate, reasonable, meaningful, fair, and relevant language production, we believe that the cognitive underpinnings of a task-based approach need to be reconsidered and enhanced through the integration of critical thinking skills. A number of studies in the literature confirm the positive relationship between language learning and critical thinking skills (Afshar, and Movassagh, 2017; Chamot, 1995; Shirkhani and Fahim, 2011; Rafi, 2010). However, the number of studies related to the effects of critical thinking skills on learners' oral proficiency is limited (Yang et al. 2013; Malmir and Shoorcheh, 2012; Sanavi and Tarighat, 2014) and there are not enough studies conducted specially to explore the impact of intellectual standards on the oral performance of learners in TBLT environments. Therefore, combining intellectual standards with the productive nature of reasoning-gap tasks and to investigate their effects on the notions of fluency, accuracy, and complexity (Skehan, 1998) is the theoretical point of departure for this experimental study. To address this problem, a task cycle model and a pre-task plan combining the principles of second language acquisition and critical thinking standards were designed. Additionally, a rubric for assessing the learners' oral performance based on the critical thinking standards was built. On the basis of these research foci, the research questions addressed in this study are:

1. To what extent do critical thinking standards have an impact on linguistic fluency, accuracy, and complexity in reasoning-gap tasks?

2. What do learners think about the critically enhanced reasoning-gap tasks and their effects on their general learning experiences?

LITERATURE REVIEW

Defining TBLT

A task is "a piece of work undertaken for oneself or for others, freely or for some reward" (Long, 1985:89). According to Prabhu (1987), a task can be defined as an activity designed for learners to reach some learning goals through using their reasoning and also to some extent it gives teachers the role of controlling and adjusting the process when it is necessary. Nunan (1989) and Willis (1996:23) emphasize the function of the tasks in providing "communicative purpose" and letting learners be part of meaning construction rather than purely focusing on form. Meaning-focused tasks within real-world contexts help students engage in meaningful activities and learning process, which generates new opportunities and experiences in interlanguage development

(Foster and Skehan, 1996). Tasks in language classrooms serve several purposes, as suggested by Pica (2008), Rahman (2010) and Willis (1996:35), they:

*   motivate learners to use the target language based on what they know in that language
*   provide learners with the opportunities of spontaneously interacting with their teachers or peers
*   let learners recognize alternative ways of expressing the same meanings
*   let learners negotiate meaning and take turns in interaction
*   help learners use language with a specific purpose and in cooperation
*   give learners chances to have interaction with longer and complete language output
*   push learners to test communication strategies
*   build learner confidence with achievable communicative goals.

Comparing a task-based approach with a Present (Perceive)-Practice-Produce (PPP) cycle, criticisms against a PPP sequence are inevitable. In the first and second stages of the PPP cycle teachers are expected to have the role of the information provider and the controller of the activity sequence and oral production and in the last stage most of the time they are the observer of students' performance (Byrne, 1976:2). However, in critical TBLT design in all stages based on the nature and demands of the given task, teachers take on the facilitator and guide roles, while learners are more in the roles of activity planner, information hunter, meaning negotiator, and active performer of linguistic and non-linguistic learning behaviors.

Types of Tasks

According to Prabhu (1987), tasks can be grouped into three types—information-gap, opinion-gap, and reasoning-gap tasks. In information-gap activities, learners are required to exchange information with their partner who does not know anything about the related information and needs to collaborate and interact with the other member to complete the task. On the other hand, opinion-gap activities are more about indicating "a personal preference, feeling, or attitude in response to a given situation" (Prabhu, 1987: 47). The last type of reasoning-gap activities involve the other cognitive skills of inference, reasoning, deriving new information from the given information. As reasoning-gap tasks are believed to encourage more collaboration and in a linguistic and cognitive sense engage learners more into task goals, they were chosen as study tasks.

Critical Thinking in Second Language Education

Critical thinking is a complex process and its sophisticated nature makes the task of defining critical thinking challenging. Therefore, it is difficult to frame a single definition. According to Ennis (1987:10), it is "reasonable, reflective thinking that is focused on deciding what to believe or do." McPeck (1981:7) sees "reflective skepticism" necessary for any incidences of activity engagement. From Green's (2005)

point of view, to be considered critical thinking, the process of "thinking" needs to be "directed toward some end or purpose." The development of essential skills alone is not sufficient for being a critical thinker. To think critically, individuals also need to possess some basic critical thinking dispositions. Some of the characteristics they should acquire can be listed as being "outcome driven, open to new ideas, flexible, willing to change, innovative, creative, analytical, communicators, assertive, persistent, caring, energetic, risk takers, knowledgeable, resourceful…" (Ignatavicius, 2001:37). Ruso (2007) suggests seven dispositions for critical thinking: truth-seeking, open-mindedness, analyticity, systematicity, critical thinking-confidence, inquisitiveness, and cognitive maturity. From these perspectives, critical thinking can be defined as an ability and inclination to properly and effectively activate all aspects of reasoning to make fair, accurate, evidential, and sensible judgments. Using reason and making good decisions is important and part of all aspects of everyday life.

Developing critical skills in educational contexts and cultivating "criticality" in students to prepare them for the global world are educational priorities. To achieve this priority, it is essential to point out that critical thinking is composed of systematicity, and this needs to be maintained with essential standards. According to Green (2005), critical thinking has these three basic features:

1. It gives shape to people's decisions and beliefs.
2. Thinking critically encourages people to engage in adequate and accurate thinking.
3. Relevant standards are naturally embedded in critical thinking.

As can be understood from these basic features, the quality of thinking and mastery of critical thinking in students can only be achieved through the internalization of some related and applicable standards. With respect to this standardization, in his definition of critical thinking, Paul also bases this "disciplined, self-directed thinking" on "a particular mode or domain of thought" (Paul, 1992: 9). Otherwise, reasoning that is not enhanced with any measures, criteria, or standards cannot go beyond the tendency of judging "inexactly, inaccurately, inappropriately, prejudicially" (Salimi et al., 2011). For Bailin et al. (1999:291), critical standards include the criteria of evaluating the acceptability of claims, the plausibility of statements, the quality of arguments, and the appropriateness of the reasons. Intellectual standards suggested by Salimi et al. (2011) are clarity, accuracy, precision, relevance, depth, breadth, logic, significance, and fairness. Clarity is based on the fact that the message needs to be conveyed. Besides having the quality of being clear, critical thought must be based on accurate information that can be supported with evidence. For the precision and relevance standards, criticality necessitates details as well as relevant facts, knowledge, and ideas. Additionally, these details and information need to provide a deeper understanding and consider other perspectives. Finally, excellence in critical thinking requires the intellectual acts of making logical, fair, and meaningful connections and judgments. In short, these standards aim to "give an end" to the students because we cannot

expect "excellence" in their thinking if we do not give them one (Salimi et al., 2011). In this study, we also tried to create a context in which students could stimulate their thinking skills in the target language through internalizing intellectual standards. Based on the standards integrated into the task model, students were first asked to gather accurate, relevant, and fair information from reliable sources. Then, students were put in a process of in-depth analysis of the collected information, considering all essential details related to the problem stated in the assigned reasoning-gap tasks. As the next standard, they were encouraged to broaden their thinking by looking at the problem from different perspectives. Finally, after analyzing the ideas, concepts, judgments, and any other kinds of representations of information concerning the task problem, students' reconstruction and reformulation of their knowledge was facilitated. In this way, to create their approach to the problem and generate their solutions, they could activate both their critical and creative thinking, which are "inseparable, integrated, and unitary" (Salimi et al., 2011).

METHOD

Research Design

The study aims to assess the impact of reasoning-gap tasks on the linguistic and cognitive development of the students. The tasks were designed following linguistic and critical thinking intellectual standards. The data were collected through the designed critical task rubric, recorded classroom interaction samples, and semi-structured interviews. Both experimental and control groups were given four reasoning-gap tasks on different topics. Only the experimental subgroups were trained on the critical thinking standards and informed about the main criteria included in the actual task checklist. Students were allowed to choose their partners to give them the right to choose the person with whom they thought they could study better.

Participants

The participants were 16 undergraduate students of English Language Teaching; 8 were female and 8 were male. Groups were randomly grouped as a control and experimental group. The experimental and the control groups were also divided into 4 subgroups and each subgroup was assigned with one of the four reasoning-gap tasks. While the control group did not receive any special training regarding critical thinking standards, the experimental group did. In these teaching sessions, the rubric was introduced and criteria on the rubric were analyzed and applied within different tasks. Students were also informed about different task types, roles, and task behaviors.

Critical Task Model and Stages

In terms of offering natural language learning and use, building a learner-centered learning environment, and providing effective communication opportunities, traditional approaches are inappropriate and insufficient (Swan, 2005).

To attain the goals of cultivating more intellectually disciplined learning, and more learner consciousness on linguistic and non-linguistic aspects of language, and to advocate the cognitive, affective, and behavioral development of each learner, tasks need to be viewed as core devices of the language teaching process. Tasks are good devices that allow learners to master the target language by activating their psychological and neurobiological mechanisms. In the second language education context, with the right opportunities, language tasks need to prepare students for life. In addition to helping learners to master the language linguistically, with decision-making and information gap activities, they also need to push learners to use their reasoning and intellectual faculties to raise their consciousness and develop awareness about critical daily issues rather than focus only on the formal properties of language. As shown in Table 1 the model used in this study was shaped around three dimensions of linguistic, cognitive, and socio-behavioral development.

The task model designed in this study had three stages. The first was the pre-task stage which provided details related to the topic, task type, and any related terms or expressions to increase learners' familiarity with the task. This stage aimed to raise consciousness, increase the readiness level of the students, and activate the students' background knowledge through elicitation techniques and critical questions. In this stage, students chose the partners or group members with whom they were going to collaborate and construct new learning experiences. In this part, students were pushed into a reflection process in which they reflected more on how they were going to plan, carry out the search, find, analyze, and evaluate information to complete the task. The second stage was the task cycle stage, with appropriate linguistic support; the teacher facilitated the students' language output through critical questions and negotiation techniques to encourage students to have accurate, fluent, and complex target language. Additionally, with cognitive assistance as a facilitator, the teacher helped students to have a critical analysis of the gathered information and the process. In this stage, based on the intellectual standards of critical thinking such as clarity, accuracy, depth, breadth, and logic, the teacher guided students to produce logical ideas and find plausible solutions to achieve the goals of the task. In the final post-task stage, learners performed their tasks to their teacher and peers and they got their linguistic and non-linguistic feedbacks. The task performances were video recorded both for student and teacher review. Based on these evaluations and their reflections on their task behaviors, learners had a chance to restructure their language output and see alternative ways of achieving the required task demands.

The linguistic dimension of the oral performance of the participants was based on the notions of fluency, accuracy, and complexity (Skehan, 1998). According to Ellis (2003: 342), fluency is "the extent to which the language produced in performing a task manifests pausing, hesitation, or reformulation." It is used to refer to the ability to use ready-made chunks of language and dealing with communication breakdown with communication strategies. Accuracy, on the other hand, is the use of target language under its norms and it was

adopted as the production of grammatically correct structures (Brumfit, 1984). Finally, the last dimension, complexity, means the production of different structures with different syntactic processing and it is about "the size, elaborateness, richness, and diversity of the learner's linguistic L2 system" (Housen and Kuiken, 2009). In this study, it was focused on finding out to what extent learners were able to use specific words, expressions, idioms, and different clauses.

The critical thinking dimension of the task model included the intellectual standards of clarity, accuracy, depth, breadth, and logic (Paul and Elder, 2009). The standard of clarity was integrated into the model to check whether the learners' statements were clear enough and free from ambiguity. Accuracy allowed us to detect whether the stated utterances were true information or not. The criterion, depth of reasoning or knowledge, defines the level of how deeply participants need to search, analyze, or understand information. It also designated the procedures that learners were supposed to apply to reasonably address the problem, its causes, effects, and thereby they could attain effective solutions and make logical decisions. The standard of breadth aimed to promote a person who had multi-perspective view and could apply that trait in any situation or conditions. Breadth of knowledge and understanding asked the participants to seek out, understand, and also learn to appreciate the other perspectives. Lastly, logic comprised the quality of reasoning that participants perform to attain sensible, valid, and consistent knowledge and deduction.

Reasoning-gap Tasks

The topics chosen for the reasoning-gap tasks were:
- **Task 1:** Separation and divorce; its causes and effects on the family members; and any possible solutions to avoid divorce and strengthen marriage.
- **Task 2:** Sharing household chores; creating a system for a fair division of domestic chores.
- **Task 3:** The problem of early teen marriage; the factors influencing teen marriage; its effects on adolescents' education; and effective ways to prevent it.
- **Task 4:** Gender inequality at the workplace; gender pay gap; its main reasons; achieving female empowerment.

Each sub-group of the control and experimental groups was assigned one of these tasks. Both non-treatment and treatment groups were verified as having completed the task but only the experimental groups were informed with required intellectual criteria and standards.

Data Collection Tools

Multiple tools, such as students' video recordings, pre-task planning report forms, task evaluation rubrics, and semi-structured interview protocols, were used to collect data. Through mixed-methods design, both quantitative and qualitative data collection tools were incorporated. The triangulation of the data with multiple methods helped us to strengthen the credibility and validity of findings and provided a comprehensive understanding of the research phenomenon (Creswell and Clark, 2011; Greene et al., 1989). As

a first data collection tool, the task groups in the experimental group were provided with a pre-task plan for their tasks (see Appendix A). The parts of the reports were divided into three main sections as daily schedule, linguistic, and intellectual dimension. The daily schedule part was designed to get enough details related to participants' role sharing, group discussions, and the activities that they did before to give a successful task performance. The linguistic dimension aimed to increase the participants' awareness of lexical and syntactic richness so that they could also focus on the lan-

Table 1. Critically Enhanced Task Model

Stages		Goal	Task Techniques
Pre-Task		Introducing the Task	Elicitation Consciousness Raising
	Linguistic	Introducing New Language	
	Critical Thinking Skills	Activating Schemata Planning	
Task		Task Cycle Task Completion	Negotiation of Meaning Information Exchanging Problem-Solving Decision-Making
	Linguistic	Accuracy-Complexity-Fluency Restructuring Language	
	Critical Thinking Skills	Search-Analysis-Evaluation-Creation Universal Intellectual Standards (clarity-accuracy-depth-breadth-logic)	
Post-Task		Task Reporting	Presenting Results Reflection on Task
	Linguistic	Teacher–Student–Peer Feedback; Error Correction	
	Critical Thinking Skills	Self-Regulation Reflecting on the Linguistic, Cognitive, and Attitude Development	

Table 2. Task Performance Rubric

Language development			
Fluency	**Excellent**	**Good**	**Needs improvement**
False starts: Starting an utterance then aborting it.			
Repetitions: Repetitions of words, phrases, or clauses			
Hesitations (Stuttering): Problems during uttering a word; uttering initial phoneme or syllable(s) then completing the word.			
Pronunciation: No mispronunciations, speaks clearly			
Accuracy			
No error in syntax (clauses and sentences)			
No error in morphology (word forms noun, verb, adj. etc.)			
No error in semantics (meaning of individual words and the given message)			
Complexity			
Word choices: Variety in topic-related word choice			
Sentence Variety: Different types of sentence (simple, complex, compound, compound/complex sentences)			
Cognitive development			
Intellectual standards			
Clear and enough illustration of the message			
Correct information supported with evidence			
Complete and detailed analysis and presentation of information			
Giving all point of views or perspectives related to the topic			
All given information is well-organized and sensible.			

*Improving EFL Learners' Oral Production through Reasoning-gap Tasks Enhanced with Critical Thinking
Standards: Developing and Implementing a Critical TBLT Model, Pre-Task Plan, and Speaking Rubric*

45

guage itself not only the task non-linguistic requirements. Lastly, the intellectual dimension encouraged learners to gather more credible, accurate, fair, and comprehensive information.

To measure the participants' speaking performance, a rubric was developed. The rubric was designed around two categories that learners were expected to achieve. The first category is related to language development as fluency, accu-racy, and complexity (Skehan, 1998). This part measures the learners' oral production from the perspectives of speak-ing "faster and with fewer instances of silence and repair," producing "more error-free utterances," and hav-ing "more complex grammatical structures and vocabu-lary" (Foster and Skehan, 1996). The second main category describes cognitive development under five sub-points of intellectual standards as clarity, accuracy, depth, breadth, and logic. Learners' conceptualization of the topic, the conclu-sions that they drew, and the solutions they produced were judged. Each descriptor in the rubric was rated with three Likert points as "Excellent", "Good", and "Needs Improve-ment"(see Table 2).

Semi-structured interviews were conducted with the par-ticipants to encourage them to openly express their opinions regarding their reasoning-gap task. The relevant points were determined and formulated with open-ended questions. The points indicated in the semi-structured interview guide were:

- reasoning-gap task enhanced with critical thinking stan-dards and its impact on language development
- reasoning-gap task and cooperative learning
- the use of video recording for task performance and its-role in the feedback section

Each interview was tape-recorded and later transcribedso as not to miss any important points that might emerge during the interview becdue to semi-structured interviews' open and flexible nature.

Data Analysis

For data-gathering, all students present both in the control and the experimental group were recorded and the recorded data were transcribed. The rubric designed for assessing the task performance was used. The results from all groups were compared based on the variables stated on the rubric. To increase the reliability of the applied rubric, one more rater was also asked to use the rubric to rate the participants' speaking performance. Based on 20 percent of the recorded data, the inter-rater reliability between two raters was very high (Cronbach's α=.98).With respect to the other statistical analyses, multivariate analysis of variance (MANOVA) was conducted to see whether there were any differences between treatment and non-treatment groups on the variables of flu-ency, accuracy, and complexity. To test the homogeneity of variance, Levene's test was used. The test results were not statistically significant, which indicated that the data had a normal distribution. The transcribed oral productions were also analyzed through the digital tool Text Inspector. Lexical diversity, the European Vocabulary Profiles of the vocabu-lary used, and the use of meta-discourse markers were the analyses carried out by the Text Inspector. In addition to

these analyses, for the measurement of the syntactic density, main and subordinate clauses were noted and the ratio of all clauses to the number of T-units was calculated for each task performance. The interview data were recorded and tran-scribed for coding.

FINDINGS

The study employed a mixed-methods research design to investigate the impact of critical thinking standards on the linguistic dimensions of students' oral performance (fluency, accuracy, and complexity).

The results of MANOVA in Table 3 indicated that the use of critical thinking standards in reasoning-gap tasks sig-nificantly influenced the mean scores of accuracy [$F (3, 11)$ = 10.12, p=.00; Partial Eta Squared =.43, indicating a large effect size] and the mean scores of complexity [$F (3, 11)$ = 17.47, p=.00; Partial Eta Squared=.57, indicating a large effect size]. However, the integration of critical standards into the task design did not have any significant effect on the linguistic variable of fluency [$F (3, 11)$ = 1.71, p=.21; Partial Eta Squared=.11].

Findings Related to Linguistic Complexity

For the difference of clausal density between two groups, the subordination Index (SI) was computed. SI is the ratio of the total number of clauses (main and subordinate) to the total number of main clauses. Table 4 indicates that, compared to the control group, participants in the experimental group had more utterances and more complex sentences with subordi-nate clauses with conjunctions and relative pronouns, which provide a smooth transition between the clauses and give more essential and more detailed information.

Table 3. Multivariate Analysis of Variance (MANOVA) between Control and Experimental Groups

Variable	*F*	Hypo df	Error df	Sig.	*Partial Eta Squared*
Fluency	1.710	3	11	.21	.116
Accuracy	10.127	3	11	.00*	.438
Complexity	17.474	3	11	.00*	.573

Table 4. Clausal density between groups

Task name		Total clauses	Main clauses	SI
Divorce	Control Gr.	24	19	1.2
	Experimental Gr.	53	41	1.3
Sharing household chores	Control Gr.	32	27	1.2
	Experimental Gr.	55	39	1.4
Early teen marriage	Control Gr.	48	32	1.5
	Experimental Gr.	77	51	1.5
Gender inequality	Control Gr.	37	26	1.4
	Experimental Gr.	65	35	1.9

In regard to the measurement of lexical variation in the oral performance of the groups, TTR was calculated. For the TTR calculations, the last one hundred words of transcribed oral texts were used with the purpose of eliminating any potential flaws related to the text. TTR comparisons between subjects indicated that all subgroups in the experimental group had a higher ratio compared with the ones in the control group (see Table 5).

In addition to the measures of TTR, which give the ratio of different types of words to the overall words in the text, the measures of vocd-D and MTLD were also calculated. McCarty and Jarvis (2010) suggest that one index is not enough to get reliable results related to the lexical diversity. Whereas vocd-D gives the "sums of probabilities" of each type based on the sample 25-50 tokens, MTLD is "the mean length of sequential word strings in a text" (McCarty and Jarvis, 2010). These two measures were carried out with a professional web tool called Text Inspector. Compared to the ones in the control group, it can be concluded that most of the sub-groups in the treatment group had higher lexical diversity in their task presentations.

The vocabularies used by the participants were cate-go-rized in accordance with the Common European Framework. The analysis of the words and the phrases used by the groups at the upper levels of B2, C1, and C2 were chosen in order to determine the effect of task preparation and intellectual criteria on the participants' vocabulary knowledge. As presented in Table 7, the treatment group showed higher performance than the non-treatment group. Some of

the vocabulary and phrases used by the experimental group at the levels of B2, C1, and C2 were *expense, strain, with crossed arms, tend to* (divorce); *hire help, renegotiate, chore charts* (sharing household chores); *dilemma, purity, morality, enforce* (early teen marriage); *appreciate, violation, workforce, persistent* (gender inequality).

Table 8 illustrates the frequencies related to the use of meta-discourse by two groups. The functions of the instances of meta-discourse (Hyland, 2004) were analyzed within the standards of critical thinking. These functions were grouped as showing evidence for the gathered information, using symbols for logically connecting clauses and ideas, using "boosters" to emphasize the given message or relational markers for building a relationship with the audience, and lastly using code glosses to elaborate what has been said. In all these categories, it was recorded that the participants in the treatment group outperformed the control group.

Findings from Semi-Structured Interviews

Semi-structured interviews with student teachers provided detailed information related to three main issues of the research foci. All teacher candidates had positive views on the impact of reasoning-gap tasks on their target language development. Students expressed that in terms of developing life skills, working with the assigned tasks allowed them to search, analyze, and synthesize the collected information based on their own life experiences and intellectual standards. They all noted that tasks also engaged them in the process of recalling the task-related vocabulary and expressions and let them learn the new ones. Additionally, they emphasized that they had the opportunity of practicing authentic target language through reading in the target language, checking pronunciation of the newly learned words, and working on more complex sentences.

Applying intellectual standards of reasoning

The given standards for thinking aimed to guide the participants to have more accurate, impartial and plausible information. They were also required to provide evidence and research-verified information during the process of searching, analyzing, evaluating, and synthesizing their solutions for the assigned problems. To ensure the reliability of the information, all student teachers stated that they mostly preferred to use educational sites and compared the information with the other sites and sources. Most of them also noted that since the topics were from real life, they went through this

Table 5. Lexical diversity

Task name		Control Gr.	Experimental Gr.
Divorce	Tokens	101	101
	Types	66	69
	TTR	65	68
Sharing household chores	Tokens	101	101
	Types	66	75
	TTR	65	74
Early teen marriage	Tokens	105	101
	Types	62	74
	TTR	59	73
Gender inequality	Tokens	105	103
	Types	58	70
	TTR	55	68

Table 6. Lexical diversity: vocd-D and MTLD

Task Name	vocd-D		MTLD	
	Control Gr.	Experimental Gr.	Control Gr.	Experimental Gr.
Divorce	79.70	91.28	49.31	49.31
Sharing household Ch.	67.75	131.08	36.28	114.25
Early teen marriage	61.71	100.98	63.06	105.79
Gender inequality	47.58	89.72	48.62	63.42

Improving EFL Learners' Oral Production through Reasoning-gap Tasks Enhanced with Critical Thinking Standards: Developing and Implementing a Critical TBLT Model, Pre-Task Plan, and Speaking Rubric

47

Table 7. English Vocabulary Profile between Subjects
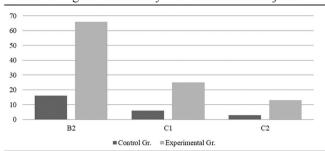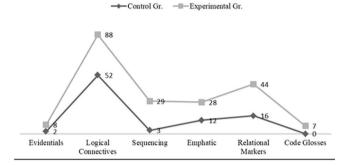

Table 8. The Use of Meta-discourse Markers between Subjects

evaluative process based on either their own or other people's life experiences. As they provide an evidence base for the accuracy of the collected data, three of the participants indicated that they mostly tried to gather statistical information. The two other teacher candidates emphasized that they filtered information culturally since they thought that some of the data in the target culture did not fit their cultural paradigm. Finally, most of the participants noted that they had already started to transfer the newly learned information to new contexts and to their life.

Critical thinking and collaborative learning

Based on the study findings, task-driven collaborative learning yielded positive effects on the students' linguistic, cognitive and personal development. Collaborative learning, as defined by Mercer and Littleton (2007), helped to engage participants "in a coordinated, continuing attempt to solve a problem or in some other way construct common knowledge". In this respect, reasoning-gap tasks within its open-ended nature provided learners with a collective sense which led them into the process of "knowledge coconstruction", shared feeling, skills, and experience (Cohen, 1994; Trentin, 2010). Using intellectual standards for collaborative learning not only enhanced the quality of individual and collective reasoning but also provided participants with a platform on which they could share their knowledge and ideas, which was more conducive to having lifelong learning. Concerning the positive impact of cooperation around a shared goal and common problem, eight out of ten participants pointed out that they had the chance to share their points of view and understand diverse perspectives. Due to having a dominating personality, three participants indicated that they thought

that they would have difficulty in working with someone else. However, at the end of the task, they agreed that sharing authority was not a threat to the control over their learning; rather, it was a facilitative process in which they could learn from others and could have someone who could compensate them for their linguistic and cognitive fallacies.

The use of video recordings for oral performance

Another important task element that was also emphasized by the participants during the feedback session was the use of the video recordings related to their actual task performance. To help the participants to recall their performance, let them have the right interpretations of the trainer's comments, and facilitate the process of making an association between the given feedback and their task behaviors, learners were asked to watch their video recordings and reflect on the given written comments. Concerning the video-assisted feedbacks, all participants pointed out that viewing their task provided them with an opportunity to engage in more critical reflection on their target language performance.

DISCUSSION

This research aimed to investigate empirically the effects of the integration of the intellectual standards into TBLT. The statistical analyses revealed significant differences between the two groups in terms of their L2 speaking performance as the experimental group outperformed the control group in their post-test performance proving that employing critical thinking skills contributed to their L2 speaking proficiency to a great extent. In this respect, the findings of this study confirm findings of the studies conducted by several researchers in the relevant field (Malmir and Shoorcheh, 2012; Sanavi and Tarighat, 2014). In their study, Malmir and Shoorcheh (2012) found that introducing critical thinking skills helped students to maintain clarity and coherency in their oral performances. Moreover, thanks to being exposed to training on critical thinking standards, students achieved more logical inferences during their oral performance. Likewise, Sanavi and Tarighat (2014) concluded that critical thinking standards were essential for promoting students' performance in L2 speech and L2 in general. Therefore, they should be integrated into EFL classes.

Considering the effect of critical thinking within the reasoning-gap tasks, the findings showed that the use of such critical thinking standards as accuracy, depth, breadth, and fairness can prompt learners to show more effort to reach correct, impartial, and multidimensional information. In addition to the efforts to achieve communicative success, the intellectual demands pushed the task performers to comprehend the situation, seek the essential information, negotiate, and reconceptualize the aspects that emerged. This intellectual process not only activated learners' reasoning faculties but also stretched their interlanguage. As noted by Ennis (1993) and Skehan (1998), reasoning-gap tasks provided learners with a context in which they focused on specific concepts, expressions, and utterances to have more fluent, accurate and complex language output. Fluency, accuracy, and

complexity were three linguistic dimensions that this study aimed to improve through reasoning-gap tasks. In regard to fluency, in addition to being accurate in terms of using structures, the speaker of that language also needs to use "the units and patterns automatically at normal conversational speed when they are needed" in order to be fluent (Hartmann and Stork, 1976: 86). Accuracy is a "learner's capacity" to deal with "interlanguage complexity" (Skehan, 1996:46). Lastly, complexity is "the stage and elaboration of the underlying interlanguage system" (Skehan, 1996:46). Based on the statistical findings and interview data, we found that the integration of the critical thinking skills into the pre-task stage encouraged learners to put all their effort into building more accurate sentences that were error-free in terms of syntactic, morphological, and lexical forms (Ahmadian and Tavakoli, 2010). As regards complexity, it was recorded that participants in the experimental group used more advanced and topic-specific vocabulary, expressions, idioms, and complex clauses with different conjunctions and meta-discourse markers in order to fulfill the intellectual standards of depth and breadth, which are the abilities of dealing with "complexities" of the problem and looking at the problem from other perspectives, respectively (Paul and Elder, 2009). In this regard, the findings of this study showed that pre-task planning enhanced with reasoning skills has a close link with accurate (Wendel, 1997) and complex production (Yuan and Ellis, 2003). Concerning the dimension of fluency, having clarity, native-like speaking, and being easily understood by the audience were the other criteria that learners needed to take into consideration for their oral performance. As for the linguistic components of accuracy and complexity, learners emphasized the positive impact of pre-task planning on their oral fluency during actual task performance, which replicated the findings of Skehan and Foster (2012). With regard to the oral performance, learners stated that during the pre-task planning stage, they had time to check the pronunciation of the words and to practice native-like speaking, paying more attention to rhythm and intonation. However, in terms of fluency, the statistical analyses did not indicate any significant difference between groups. This showed that while critical thinking standards had more instant and positive effects on students' accurate and complex language production, they did not have any remarkable effect on their fluency.

Another important conclusion that we can draw based on our investigation is that there is a positive mutual relationship between critical thinking and cooperative learning. Learners stated that being required to apply certain standards to the process of information gathering, evaluating, and reporting helped them to discuss and dive deeper into the problem with less manipulation of personal biases or untrustworthy sources. On the other hand, they indicated that working on a common problem and trying to find potential solutions to it in a group allowed them to do more careful and critical planning and most importantly let them see things from different viewpoints. They also noted that, through collaborating with other students, the massive and subjective process of making logical and evidential arguments became less challenging. In this regard, several studies (Cooper, 1995; Gokhale,

1995; Johnson and Johnson, 1986; Olivares, 2005) support the claim that cooperative learning has an important role in improving learners' critical thinking skills. Overall, the findings from this study are significant and essential for further studies and classroom practices as they point out the fact that critical thinking standards can play a key role in helping learners to master the foreign language through contributing to their awareness and action on producing more fluent, accurate, and complex language.

CONCLUSION

This research study investigates learners' oral proficiency and their perceptions related to the integration of critical thinking standards into TBLT. In this study, intellectual standards were applied in a designed TBLT model in order to foster learners' intellectual engagement before and during performing reasoning-gap tasks. A mixed-methods research design was employed and the data collection process was carried out through video recordings of the learners' oral performances and with semi-structured interview protocols. The data related to the oral performances were transcribed and analyzed through the use of specifically designed speaking rubrics and a web-based text analysis tool called TextInspector. The interview data were also transcribed and put into the process of content analysis. In conclusion, the findings indicated that applying intellectual standards to the reasoning-gap task had positive effects on the learners' oral performance. Based on the study results, the participants in the training group which was explicitly trained on critical thinking standards and exposed to these standards both during pre-task and during task stages showed successful oral performance with respect to three linguistic dimensions of fluency, accuracy, and complexity. Moreover, regarding accuracy and complexity aspects, the statistical analysis revealed that there were significant differences between the control and experimental group. Particularly related to the syntactic, morphological, and semantic analysis, learners in the experimental group displayed higher accuracy in their task performance. Again the same group was indicated significantly more successful in achieving syntactic and lexical complexity in their speech.

Regarding the learners' responses to the interview questions, all participants in the treatment group agreed on the positive effects of critical thinking standards on their oral performances. Most learners also pointed out that applying intellectual standards to their information search and use process helped them to reach more accurate, reliable, and fair information related to their reasoning-gap task topics. Developing learners' oral proficiency is not an easy task. Learners need to devote time and energy to practice and seek ways to develop their oral skills in the target language. In this difficult process, teachers should create optimal conditions for learners to develop their speaking skills. The use of critical thinking standards within the context of TBLT is one of these options. It is also essential to keep in mind that students may show resistance to thinking critically or simply they may not know how to think critically. For this reason, teachers should first believe in the importance of teaching

critical thinking skills and then show commitment and consistency in integrating these skills into their teaching. In this way, critical thinking skills could be a natural component of the language learning process.

REFERENCES

Afshar, H.S., and Movassagh, H. (2017). On the relationship among critical thinking, language learning strategy use and university achievement of Iranian English as a foreign language majors, *The Language Learning Journal*, 45, 382-398, DOI: 10.1080/09571736.2014.914238

Ahmadian, M. J., & Tavakili, M. (2010). The effects of simultaneous use of careful online planning and task repetition on accuracy, fluency, and complexity of EFL learners' oral production. *Language Teaching Research,* 15(1), 35-59.

Bailin, S., Case, R., Coombs, J. R., & Daniels, L. B. (1999). Conceptualizing critical thinking. *Journal of Curriculum Studies*, 31(3), 285–302.

Bruton, A. (2005). Task-based language teaching: For the state secondary FL classroom?, *Language Learning Journal,* 31:1, 55-68.

Brumfit, C. (1984). *Communicative Methodology in. Language Teaching: The Roles of Fluency and. Accuracy.* Cambridge: Cambridge University Press.

Byrne, D. (1976). *Teaching Oral English.* London: Longman.

Chamot, A. (1995). Creating a community of thinkers in the ESL/EFL classroom. *TESOL Matters,* 5(5), 1-16.

Cohen, E. G. (1994). Restructuring the Classroom Conditions for Productive Small Groups. *Review of Educational Research,* 64, 1-35.

Cooper, J. L. (1995). Cooperative learning and critical thinking. *Teaching of Psychology,* 22(1), 7-9.

Creswell, J. W. & Clark, V.L.P. (2007). *Designing and Conducting Mixed Methods Research.* Thousand Oaks, CA: Sage.

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

Ellis, R. (2009). Task-based language teaching: sorting out the misunderstandings. *International Journal of Applied Linguistics,* 19(3), 221-246.

Ennis, R.H. (1987). A taxonomy of critical thinking dispositions and abilities. In J. Boykoff Barron & R. S. Steinberg (Eds.), *Teaching thinking skills: Theory and practice* (pp. 11 – 26). New York: Freeman and Company

Ennis, R. H. (1993). Critical Thinking Assessment. *Theory into Practice,* 32, 179-186.

Foster, P. and Skehan, P. (1996). The influence of planning on performance in task-based learning. *Studies in Second Language Acquisition,* 18: 299-324.

Gokhale, A. A. (1995). Collaborative Learning Enhances Critical Thinking. *Journal of Technology Education,* 7, 22-30.

Green, C. (2005). Integrating extensive reading in the task-based curriculum. *ELT Journal,* 59(4), 306-311.

Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). Toward a Conceptual Framework for Mixed-method Evaluation Designs. *Educational Evaluation and Policy Analysis,* Vol 11, No. 3, pp 255-274.

Hartmann, R. R. K. & Stork, F. C. (1976). *Dictionary of Language and Linguistics.* New York: Wiley.

Housen, A., & Kuiken, F. (2009). Complexity, accuracy, and fluency in Second Language Acquisition. *Applied Linguistics,* 30(4), 461-473.

Hyland, K. (2004). Disciplinary interactions: Metadiscourse in L2 postgraduate writing. *Journal of Second Language Writing.* Vol. 13, 133–151.

Ignatavicius, D. (2001). Critical thinking skills for at-the-bedside success. *Nursing Management, 32*(1): 37-39.

Johnson, R. T., & Johnson, D. W. (1986). Cooperative learning in the science classroom. *Science and children,* 24, 31-32.

Long, M. (1985). A role for instruction in second language acquisition: task-based language teaching. In K. Hyltenstam and M. Pienemann (eds.), *Modelling and assessing second language acquisition.* (pp. 77-79). Clevedon: Multilingual Matters.

Malmir, A., & Shoorcheh, S. (2012). An investigation of the impact of teaching critical thinking on Iranian EFL learners' speaking skill. *Journal of Language Teaching and Research,* 40(3), 608-617.

McCarthy, P.M., & Jarvis, S. (2010). MTLD, vocd-D, and HD-D: A validation study of sophisticated approaches to lexical diversity assessment. *Behavior Research Methods,* 42, 381392. doi: 10.3758/BRM.42.2.381

Mcpeck, J. E. (1981). *Critical Thinking and Education.* Oxford: Martin Robertson.

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

Mercer, N., & Littleton, K. (2007). *Dialogue and the development of children's thinking.* London: Routledge.

Olivares, O. J. (2005). Collaborative critical thinking: Conceptualizing and defining a new construct from known constructs. *Issues in Educational Research,* 15.

Paul, R. W. (1992). Critical thinking: What, why, and how? *New Directions for Community Colleges,* 1992(77), 3–24.

Paul, R. W. & Elder, L. (2009). *The miniature guide to critical thinking concepts & tools (6th ed.).* Dillon Beach, CA: The Foundation for Critical Thinking.

Pica, T. (2008). Task-based instruction. In N. Van Deusen-Scholl & N. H. Hornberger (Eds.), *Encyclopedia of language and education* (2nd ed., Vol. 4, pp. 71–82). New York, NY: Springer Science+Business Media, LLC.

Prabhu, N .S. (1987). *Second language pedagogy.* Oxford: Oxford University Press.

Rafi, M. S. (2010). Promoting critical pedagogy in language education. *International Research Journal of Arts & Humanities,* 37, 63-73

Rahman, M. M. (2010).Teaching Oral Communication Skills: A Task-based Approach. *ESP World,* 9 (1)

Ruso, N. (2007). The influence of task-based learning on EFL classrooms. *Asian EFL Journal,* 18(2)

Salimi, A., Dadaspour, S., & Asadollahfam, H. (2011). The effect of task complexity on EFL learners' written per-

formance. *Procedia Social and Behavioral Sciences,* 29, 1390-1399. doi:10.1016/j.sbspro.2011.11.378

Sanavi, R., Tarighat, S. (2014). Critical thinking and speaking proficiency: A mixed method. *Theory and Practice in Language Studies,* 4(1), 79-87.

Shehadeh, A. (2005). Task-based Language Learning and Teaching: Theories and Applications. In: Edwards C., Willis J. (eds) *Teachers Exploring Tasks in English Language Teaching.* London: Palgrave Macmillan

Shirkhani, S., & Fahim, M. (2011). Enhancing critical thinking in foreign language learners. *Procedia-Social and Behavioral Sciences,* 29, 111-115. http://dx.doi.org/10.1016/j.sbspro.2011.11.214

Skehan, P. (1996). A framework for the implementation of task-based instruction. *Applied Linguistics,* 17: 38-62.

Skehan, P. (1998). *A Cognitive Approach to Language Learning.* Oxford: Oxford University Press.

Skehan, P., Foster, P. (2012). Complexity, accuracy, fluency and lexis in task-based performance: A meta-analysis of the Ealing Research. In Housen, A., Kuiken, F., Vedder, I. (Eds.), *Dimensions of L2 performance and proficiency: Investigating complexity, accuracy and fluency in SLA.* Amsterdam: John Benjamins.

Swan, M. (2005). Legislation by Hypothesis: The Case of Task-Based Instruction. *Applied Linguistics,* 26 (3), 376–401.

Trentin, G. (2010). *Network Collaborative Learning.* Oxford: Chandos Publishing.

Wendel, J. (1997). Planning and second language narrative production. Ph.D. dissertation, Temple University.

Willis, J. (1996). *A framework for task-based learning.* Harlow: Longman.

Yang, Y. T. C., Chuang, Y. C., Li, L. Y., & Tseng, S. S. (2013). A blended learning environment for individualized English listening and speaking integrating critical thinking. *Computers & Education,* 63, 285-305

Yuan, F. and R. Ellis. (2003). The effects of pre-task planning and on-line planning on fluency, complexity and accuracy in monologic L2 oral production. *Applied Linguistics,* 24(1).

APPENDICES

Appendix A: Pre-Task Plan

Daily schedule	Task name: Task group members
	Task-related decisions & Pre-task behaviors
Monday	
Tuesday	
Wednesday	
Thursday	
Friday	
Saturday	
Sunday	
Linguistic dfimension	
Lexical diversity	Use of topic-related, specific words, expressions, idioms; give some examples
Syntactic complexity	Building compound and complex clauses, use of conjunctions, linking words, relative clauses; give some examples
Intellectual dimension	
Accuracy	Correctness of information; use of reliable sources, cross-checking information
Depth	All complexities related to the task problem; the causes and effects of it; the possible solutions to the problem
Breadth	Other points of view; perspectives for finding alternative solutions
Logic	Is all provided information logical, related, and supported by evidence?