

ARTICLE

Task-Based Language Assessment: A Case Study in Effectiveness in EFL Contexts

Abdullah Alshakhi , Marwan Albalawi 

English Language Institute, King Abdulaziz University, Jeddah 21589, Saudi Arabia

ABSTRACT

Language assessment bears profound influence on language learning, especially in foreign language teaching and learning environments. Formative assessment using feedback and correctional strategies facilitates learning very effectively. Task-based language assessment (henceforth TBLA) employed as a formative assessment technique is a comparatively new assessment model. The present paper discusses the results obtained from a qualitative study conducted with undergraduate students at a Saudi university investigating the effectiveness of TBLA as an assessment technique as well as a learning aid helping learners improve their writing skills. Observation, semi-structured interview, and document analysis were used as data collection instruments in the study. The findings show that TBLA aligns very strongly with the given syllabus objectives and enhances learners' language use and higher order thinking skills. The findings of the present study have serious pedagogical implications in Saudi Arabian EFL contexts as the technique is found to be effective in improving learners' English writing skills.

Keywords: Task-Based Language Teaching; TBLA; Formative Assessment; Writing Skills; English as a Foreign Language

*CORRESPONDING AUTHOR:

Abdullah Alshakhi, English Language Institute, King Abdulaziz University, Jeddah 21589, Saudi Arabia;
Email: amalshaikh1@kau.edu.sa

ARTICLE INFO

Received: 22 October 2024 | Revised: 2 November 2024 | Accepted: 15 November 2024 | Published Online: 12 December 2024
DOI: <https://doi.org/10.30564/fls.v6i6.7515>

CITATION

Alshakhi, A., Albalawi, M., 2024. Task-Based Language Assessment: A Case Study in Effectiveness in EFL Contexts. *Forum for Linguistic Studies*. 6(6): 929–946. DOI: <https://doi.org/10.30564/fls.v6i6.7515>

COPYRIGHT

Copyright © 2024 by the author(s). Published by Bilingual Publishing Co. This is an open access article under the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) License (<https://creativecommons.org/licenses/by-nc/4.0/>).

1. Introduction

Assessment has a profound effect on teaching and learning^[1]. The importance of assessment in teaching and learning cannot be disregarded, as it provides teachers with the information needed to assess and review their strategies. In addition, assessment provides its users with the information needed to make high-stakes decisions that can profoundly affect the test-taker's life^[2]. TBLA is a promising method that can improve assessment practices in the Saudi English as a Foreign Language (henceforth EFL) contexts by serving as an assessment technique to test learners' English skills and by introducing authenticity to the assessment process^[3]. However, in Saudi Arabia research on TBLA is still in its nascent state, and, therefore, further research in this academic field is a requirement of the time. A preliminary review of existing literature on the subject revealed that there is very little research work on TBLA, especially in Saudi Arabian university EFL contexts, though in a number of classrooms task-based language teaching approach is followed to teach English. Formative assessment also takes place, but it is more often based on communicative language teaching/testing model^[4-5]. The researchers encouraged a few teachers to implement task-based formative assessment model to assess learners' English writing skills. The practice was aimed at assessing students' English writing skills in the Saudi context, to gain qualitative insights into how TBLA affects EFL learners' writing skills. At a later stage, the researchers evaluated the effectiveness of TBLA as an assessment method. The findings of the study may prove useful for language teachers in Saudi Arabia for a better understanding of the pedagogical and practical consequences of adopting TBLA in the given context.

2. Research Problem

Task-based language teaching has been in practice in EFL environments for a long time now^[6]. However, TBLA is not widely practiced as scholars report difficulties with performance-based assessment. So, little is known on the effectiveness (or otherwise) of TBLA technique in EFL environment since, even if English language is taught in some classrooms following task-based language teaching model, for assessment purposes the teachers still follow ei-

ther traditional assessment methods or apply communicative language teaching approach-based model of assessment. The existing research literature on the subject is also of little help as research on the topic is scanty or altogether missing in EFL teaching contexts like Saudi Arabia.

3. Literature Review

3.1. Task-Based Language Assessment

TBLA is an assessment method that shifts the focus from what the learner knows to what the learner can do^[7]. This "can-do" assessment, which uses tasks to elicit test-takers' performance, has been identified by different names by researchers in the assessment field. For instance, Brindley^[8] called it "task-centered assessment" (TCA) and defined it as "the process of evaluating, in relation to a set of explicitly stated criteria, the quality of the communicative performances elicited from learners as part of goal-directed, meaning-focused language use requiring the integration of skills and knowledge" (p. 74). TBLA is not an assessment method that is used to classify test-takers into broadly defined degrees of language ability by measuring their linguistic knowledge^[9]. Instead, TBLA is used to assert the test-taker's ability to utilize their second language to complete assigned tasks^[9-10]. In fact, Long and Norris^[9] stated that "genuinely TBLA takes the task itself as the fundamental unit of analysis, motivating item selection, test instrument construction, and the rating of task performance" (p. 600). TBLA utilizes tasks as a central component for stimulating and observing language used for an authentic purpose to understand what test-takers are capable of doing with their language abilities^[11-12].

3.2. The State of TBLA Literature

The subject of TBLA has been under discussion among researchers for roughly two decades now^[8,10-14]. Early studies^[13-15] were concerned with issues related to the task-design and measurement aspects of the assessment and suggested ways in which TBLA's generalizability could be improved. Some scholars^[8,10,11] focused on the difficulties of performance evaluation and discussed practical issues that TBLA users may confront. Other research-

ers discussed the implementation of TBLA in different contexts around the globe. In America, Byrnes^[16] reported the role of TBLA in an adult German as a foreign language program to assess students' writing skills. Byrnes^[16] focused on a department that had transitioned from using a form-based normative approach to implementing instructions that utilized language use and meaning-focused approaches in their undergraduate curricula. Byrnes^[16] concluded that TBLA contributed to a positive washback that led to effectively linking learning goals with the desired instructional outcomes. Similarly, Ke^[17] analyzed the implementation of a formative TBLA in a Chinese language program at the University of Iowa. Their implementation of TBLA was criterion-referenced that used an analytic scoring method. They reported that TBLA enabled the teacher to adjust their teaching and provide learners with feedback regarding their development.

Byrnes^[16] and Ke^[17] showcased the benefits of using a formative TBLA. Those assessments were tailored to fit the needs of those specific contexts, which do not reflect the interactions of the assessment with other curriculum objectives, as the reported results in both studies were specific to their context and cannot be generalized. Finally, both of these studies involved the implementation of TBLA to assess languages other than the English language, which leaves a clear gap in understanding the implications of TBLA for assessing English. When it comes to the EFL context, studies such as Sarıgöz and Fişne^[18], in the Turkish context, investigated the implementation of the Integrated Language Program for Young Learners (ILTPY), a program constructed within the TBLA framework. The researchers found that ILTPY positively affected the language learning process causing improvement in learners' attainment and achievement. In addition, the program's implementation increased learners' motivation and participation, consequently increasing the EFL teachers' motivation.

The results obtained by Sarıgöz and Fişne^[18] were based on the implementation of TBLA in young learners' classrooms, which is substantially different from an adult classroom. In fact, TBLA not only taps into a test-taker's ability to use the language—it also relies on the test-taker's cognitive abilities^[19]. This limits the studies' results to

young learners and prevents their generalization to adult learners. Additionally, in Sarıgöz and Fişne's study^[18], since the learners were level A1 on the CEFR, the curriculum goals were focused on listening and speaking, with minimal reading and writing. Therefore, whatever results the researchers found were limited to those two skills. In China, Liu^[20] explored the implementation of TBLA in the Chinese context. The implementation was examined among adult EFL learners between the ages of 19 and 21. According to Liu^[20], in English writing classes in their context, students' writing is assessed in a way that they deem imperfect. Liu^[20] concluded that TBLA positively affected the students' writing. It increased the students' motivation, autonomy, and independence in their learning. TBLA is an effective way of providing feedback by serving as a tool that allows teachers to check the effectiveness of their teaching and understand the student's strengths and weaknesses.

The researcher in the study reviewed above was the one teaching the participants, thus making him an excellent source of information about the students' behavior in class because he knew them better. Thus, results, such as the learners' writing improvement, were based on quantitative evidence (in this case, the pre-test and post-test results). However, the results such as the learners' increased autonomy, enhanced self-confidence, and increased interest in writing were based on the researchers' subjective observations. The tasks used were exclusively writing-based and did not include reading components. Additionally, a holistic rating scale was used to evaluate the students' writing, raising important questions about the potential benefits or challenges that might arise if a different type of rating scale were used.

On a similar note, Chen and Wang^[21] in China examined aspects of learners' competence that were affected after more than a year of implementing a task-based language teaching and assessment framework. They also analyzed the types of tasks that have a positive impact on the language learning process and the learners' opinions of the assessment framework. The researchers found that multiple non-linguistic factors, such as organization, critical thinking, and learners' attitudes toward learning English, were positively affected by the teaching and assessment

framework. Learners' abilities, such as problem solving and their ability to evaluate, also improved. However, the results of the study were based on TBLT and TBLA in combination. Therefore, it is unclear to what degree TBLA contributed to the reported results, leaving a gap in understanding the implications of adopting TBLA with other teaching methods. Another limitation of the study was that female participants were more in number than male participants (5 men and 26 women).

In Saudi Arabian EFL contexts, Almosa^[22] analyzed the assessment practices of 22 public universities in Saudi Arabia and discovered that the concepts of summative assessment heavily impacted the universities' assessment practices. Many Saudi universities use written examinations to measure students' writing because it is a practical and fair method. Almosa^[22] concluded that Saudi universities should reconsider their assessment practices to include assessment methods that help tap into learners' critical thinking abilities.

3.3. The Research Gap

Although the task-based framework has been a topic of interest for many researchers in the Saudi context, their focus has been on task-based language teaching (TBLT), while the TBLA is largely neglected. The result is that there is no literature available on the impacts of TBLA on English language teaching and learning in Saudi contexts despite TBLT being there in practice. A cursory view at the review of existing literature also shows that there exists a research gap, particularly in Saudi Arabian post-secondary EFL teaching contexts, and the present study is a modest attempt to fill that gap in research and reference literature. The present study aims to address these gaps by implementing TBLA as an assessment for learning, using integrated tasks and a checklist-based rating scale to assess adult students' English writing skills in the Saudi context. It seeks to provide detailed qualitative insights into how TBLA affects EFL learners' writing skills and to evaluate its effectiveness as an assessment method. To better understand the role of TBLA and its impact in practical educational environments, this research will focus on the use of integrated tasks within a classroom setting as opposed to

examining its role within a specially designed educational program or curriculum that is entirely centered on TBLA.

4. Research Questions

This qualitative study, that follows the case-study approach, was designed to answer the following questions:

- RQ 1: How effective is TBLA in measuring students' writing skills?
- RQ 2: How does TBLA affect students' learning of writing?

5. Methods

In the present study, qualitative research methodology has been employed to investigate the implementation of TBLA as a formative assessment and an assessment for learning among university-level EFL learners in Saudi Arabia. The heavy reliance on qualitative method owes to the nature of the study, i.e. opinionated research. The data for the study have been collected using qualitative methods, such as observation, structured interview, and document analysis. The collected data were interpreted qualitatively to present the findings in narrative format for easy understanding.

5.1. Participants

The study participants were selected using purposeful sampling method, which is a suitable method for case studies^[23]. The participants included 26 male Saudi EFL university-level learners in the age range of 18 and 19 years. The study sample included a native EFL teacher with several years of experience teaching EFL who was assigned to teach the participants writing skills. Because the study included two TBLA implementations, during the first implementation, 22 students were present, and during the second implementation, only 14 students were present out of the total participating students. The participants had a minimum of six years of exposure to English before joining the university and their English proficiency level at the time of participation in the study was B1.

5.2. Research Context

In this study, the English course was intended for

students who enroll in majors related to science, where the main language of instruction is English. English in this course is taught for academic purposes, and it targets students' listening, speaking, reading, and writing skills. Listening and speaking are taught separately from reading and writing during the week. Regarding the assessment aspect of the course, writing is assessed using three separate marked online assignments. Students' writing is also assessed by a writing exam given at the end of the course as summative assessment. The target proficiency of the course is B1-level on the Common European Framework of Reference (CEFR). The study was conducted during the final trimester of the academic year 2022–2023, and students had progressed from earlier levels to reach this final level.

5.3. Research Design

The present study has used qualitative research design to answer the research questions since the design is suitable to studies in which participants' beliefs, interactions, attitudes, and behaviors are involved [24]. The researchers have interpreted the results obtained through data collection. The interpretation makes the researchers an essential contributing factor to the study's findings [23]. Considering the research questions, case-study approach was found suitable since it is a qualitative approach that is used to investigate a case or multiple cases for a period of time using detailed data-collection methods, such as interviews, observations, and document analysis [25]. The researchers' role in this study was that of outsiders, which means that the researchers were not a part of the group under study [26].

5.4. Task Design and Performance Rating

TBLA was used as a formative classroom-based assessment technique and as an assessment technique for learning. The assessment was expected to provide information about what the learners can do with the language they have learned, in order to identify students' learning gaps and thus improve their performance by utilizing the information obtained from the assessment in the feedback process. The target language use tasks were designed in accordance with the syllabus contents. Additionally, the

assessment was criterion-referenced, which Brown and Hudson [27] define as a test developed to elicit test-takers' performance to compare it against a set of goals or objectives. In the present study, those goals and objectives were predetermined by each unit in the book. Thus, when developing the checklist, those objectives were dealt with as the criteria of performance, making the assessment criterion-referenced.

Learners were given two tasks: the first one after they finished Unit 2, and the second one after they completed Unit 4. The tasks were integrated—they required the use of both reading and writing abilities. However, the assessment focused on the aspects of writing skill learned in those units. Those aspects of writing were translated into a set of criteria that were included in the checklist. Each task given to learners corresponded to the unit's theme. However, the assessment checklist did not exclude aspects of performance targeted by the previous units. Another checklist that took into account task fulfilment, content, organization, and language structures was used to assess the learner's performance.

The performance rating included group peer/self-assessment to elevate the formative value of the assessment. Peer assessment was used not only for feedback purposes but also as a way to elevate the authenticity of the task, where students function as an audience reviewing their peers' work. Thus, it is emphasized that language is a form of communication that is used to deliver a message rather than just comprising linguistic and grammatical structures. On a separate day following the implementation of each task, a teacher-to-student feedback session was held. During the feedback session, the teacher used the checklist to provide each student with feedback on their task performance. **Figure 1** below showcases the implementation of TBLA.

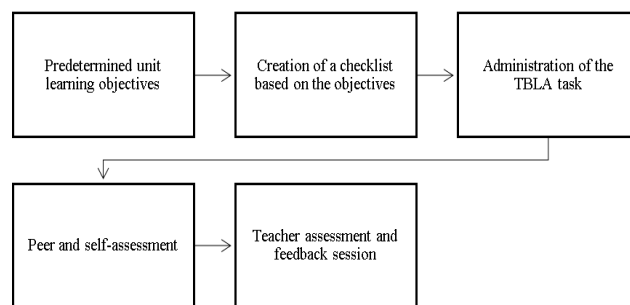


Figure 1. TBLA Implementation Process.

5.5. Data Collection Instruments

The study relied on multiple data-collection instruments. The reason for this was that relying on one data collection source instrument is insufficient to confer a deep enough understanding of the issue under study^[25]. To ensure the credibility of the findings, triangulation method was used. Triangulation method is used to improve the validity of the findings by using different sources of data for the purpose of forming themes^[28]. In this study, triangulation was affected by analyzing data from the following sources:

5.5.1. Observation

While conducting observations, the researcher was a non-participant observer, just observing without any involvement in the situation being studied^[23]. During each of the TBLA implementations, the researcher took descriptive notes on the task-administration process, learners, and the teacher's attitude during the TBLA process. The researcher also attended the teacher-to-student feedback sessions that followed each TBLA implementation, and conducted observation, taking similar descriptive notes.

5.5.2. Interview

A semi-structured interview with the teacher was conducted after the implementation of the two TBLA tasks. The interview was conducted in the teacher's office, which lasted approximately two hours. The teacher was asked predetermined questions following the interview guide. The researcher also asked follow-up questions, when needed, to gain a better understanding of the teacher's answers. If the teacher addressed a predetermined question naturally without being asked, the researcher would follow up rather than repeat the same question. The entire interview was audio recorded using smartphone audio recorder application, with the teacher's consent, to ensure accurate transcription and analysis.

The interview questions (Interview Guide) were as follows:

- How effective is TBLA in measuring students' writing skills?
- How well did TBLA align with the language

learning objectives?

- How did TBLA help measure student performance and progress?
- How reliable were the TBLA results?
- How well did the TBLA results correlate with other assessments of student language proficiency?
- What were the strengths and weaknesses of TBLA as an assessment tool?
- How does TBLA affect students' learning of writing?
- To what extent did TBLA affect students' motivation and engagement?
- How did TBLA impact students' learning of writing?
- How was the quality of the feedback and guidance provided to students based on TBLA results?
- How did students perceive TBLA as a form of assessment?
- How did TBLA affect their learning experience?
- Were there any benefits and/or challenges of using TBLA for student learning outcomes?

5.5.3. Document Analysis

Document analysis is a qualitative research instrument through which a systematic procedure is used to analyze documents to answer particular research questions^[29]. Additionally, document analysis in qualitative research is used in combination with other methods, such as interviews and observations, to triangulate data and eliminate any biases that might occur^[30]. In the present study, students' writing on the TBLA tasks given to them and their peer assessments were analyzed to gain insight into how their writing skills had been affected by the implementation. However, only the papers of 10 students were analyzed because they were the only students who took both tasks. The criteria used to identify improvements in students' performances were as follows:

- Fulfilment of Task Requirements: This criterion addresses learners' fulfilment of the task requirements by providing a complete response to the task.
- Organization and Coherence: This criterion involves the logical flow and organization of learner's thoughts, including the quality of their topic sentence, supporting details, and concluding sentence.

- **Language:** This criterion addresses learners' vocabulary, grammar, punctuation, and sentence structure.
- **Peer-Assessment Accuracy:** This criterion involves the overall accuracy of students' peer assessment of their performances on the first and second tasks.

6. Data Processing and Analysis

To analyze the collected data, thematic analysis was used to identify patterns in the collected data^[31]. After data collection, the audio recording of the interview with the teacher was transcribed. The observation notes were transferred to an observation report after each task implementation. Further, the students' writing and peer assessments on the tasks were collected and analyzed. The processed data were then coded. Later, the ZPD was used as theoretical framework to explain the findings and their implications.

7. Results

The results obtained from data analysis are presented below:

To the question on the assessment's alignment with the course goals and objectives, the teacher's response in the interview, was:

"I found it to be an accurate reflection of what's presented in the book. I mean, whatever the learning outcomes of the book would be, the assessment accurately reflected that."

The teacher added:

"It's like we are trying to make them stop, look, and think about something that they're taking at a fast pace. So, the best thing about it is that it's not extra. It's not like something that's additional that they have to do; it's something that echoes what they're doing. It's echoing what the book is saying and it's another way of honing it in so that they actually get it."

7.1. Positive Washback on Teaching

The assessment also influenced the teacher, as he was able to find the gaps in learners' knowledge and prioritize what should be learned, "Because it's another way to cor-

ner me into reinforcing what the learning outcomes are after." Regarding the assessment process, the teacher also stated: "I had the opportunity to gauge the fruits of my labor."

7.2. Focused Teaching

The teacher stated the following:

"I would make sure, obviously, to prepare my students to include all the points that we're checking for in the checklist. So, it's a recap of the important points for me. It's also a way for me to gauge, to check in which parts the gaps are. So, I can sort of cross out the points that I covered and focus on the ones that I didn't, and I can also prioritize because I can look at them and see which one of them they are not prepared for."

The teacher was also able to hold a feedback session, where he could give each of the students, who had participated in the task, personalized feedback about their performance on the task using the checklist. This allowed the teacher to provide feedback on specific parts that had contributed to the overall performance of the student.

7.3. Reinforcing the Learning Objectives

Before the assessment process, the teacher introduced the checklist to the learners. While doing so, he reinforced key aspects such as using a topic sentence, a supporting sentence, and what constitutes a response paragraph. During the feedback sessions, the teacher explained the mistakes made by the learners in their task performance.

7.4. Exposing Actual Performance

The teacher reported during the interview that throughout the assessment he had been able to capture the students' performance better than the blackboard assignments given to students. As he observed, the learners' performance on the blackboard assignments can sometimes be misleading because some students tend to use the help of others to produce the writing they submit. He also stated, regarding TBLA:

"I found, basically, that I was able to put in a corner, or catch them, or find the weaknesses of many students

who I thought were good writers. So, it's a scale that works. It doesn't come back with no readings. So, it actually comes back with a reading."

7.5. Personalized Feedback

During the feedback, learners were also able to listen to the feedback given to their peers. Furthermore, when the teacher noticed recurring issues among multiple students, he made a point to bring these to the attention of the entire class, ensuring widespread understanding. An example of this was when several students faced difficulties in fulfilling the task requirements. The teacher brought the matter to the attention of the entire class, assuring widespread comprehension. The teacher used this opportunity to emphasize the necessity of task fulfillment as well as the use of correct language, addressing a common issue and reinforcing important learning objectives.

7.6. Positive Washback on Students' Learning

The teacher expressed his happiness during the interview about the fact that the assessment was a way of guiding the learners to what he wanted them to learn:

"That's why I like this. Because it's like telling them, see this here? If you do this, I'll be happy. So, it's a way for them to do everything I want them to do or everything I want them to achieve."

The teacher also noted that the task was a factor in leading the students to what they should be learning: "The task sheet is the driving directions, but they've got the wheel." In another statement regarding TBLA, the teacher observed the following:

"This is exactly what I want. I want them to be aware of this. This is basically baby food, a rubric baby food. It's like the rubric is the actual dense meal that babies can't consume, and you're giving them; you're turning it into baby food."

7.7. Students Find Their Knowledge Gaps

During the interview, the teacher stated that students had been able to identify issues in their writing. About the benefits of the assessment, the teacher responded, "You're

able to zero in on issues that you have with your writing and be aware of them for the future. You are guided to notice certain things that you might not have known are important."

During the observation, some students responded as follows when asked about what the assessment had offered them:

- While assessing, there were mistakes I did not know about, and my peers were able to correct them, and I also got to correct my peers' mistakes.
- I got to know my mistakes.
- We got to assess our mistakes in a fair way.

7.8. Better Motivation and Engagement

On being asked how much he believed the assessment had impacted the students' motivation during its implementation, the teacher responded, "During the task, they were more motivated, for sure." He also pointed out that students showed more involvement:

"I see them more involved in writing. It's different from the quick, you know, three minutes that we give them to write something down; for example, they were able to take their time, actually be left alone, and you know, to write. So, I've noticed that they're involved. They're involved, and they're being left alone to do what they want to do. So, I didn't have to egg them on."

Regarding the learners' interaction during the task, the teacher added: "They were more at ease because they were not pressured, and they were more focused. They weren't chatting about other things because there was a task at hand."

During the peer-assessment process, increased motivation was observed; in fact, even the students who had been quiet during the discussion of reading text became more involved. The researcher also noticed that students seemed happy and motivated to assess each other's work.

7.9. Students Have More Freedom

The teacher was asked if students had more control over their writing, and he responded, "I would say they were given; they were in the driver's seat in comparison, not ideally, but in comparison." When the teacher was

asked to elaborate, he responded as follows:

“In comparison to what they’re doing, which is like, oh, all right, do page six or exercise six, work with your partner, and I’ll give you two minutes, and they have to do whatever, right? So, they’re just coming up with an answer. It’s either A or B or C, or it’s this word or this word or this word.”

The teacher added, “Here they’re being left alone, so they’re in control of the process from beginning to end. They’re in control of the process of writing from the start to the finish line.”

7.10. Student-Centered

During the assessment, more student-to-student interactions were observed. Those interactions were prompted by the requirements of the task and contained discussions regarding the reading text. For instance, on the first task, students discussed with each other the causes and effects of melting glaciers. Those interactions were heightened when students who still did not understand some of the assessment criteria began asking their peers about the meaning of it. In addition to the feedback, they provided their peers with all those interactions that were prompted by the assessment task, creating a classroom environment where the teacher’s role was minimal and limited to a facilitative role. The teacher was responsible for setting up the task, briefing the students, and answering their questions. The teacher managed disruptions caused by late-arriving students by separately explaining the task to them. Of course, not all students were interactive, as a few students opted to not participate in student-to-student interactions even though they were encouraged by the teacher to do so. However, those students began to interact with one another during the peer assessment.

7.11. Learners as Givers and Receivers of Feedback

During the interview, the teacher noted that the assessment had allowed the learners to give feedback and to receive it, as follows:

“The benefits of the assessment, I mean, you’re able to compare your writing to your partner. You’re able to

give feedback on the things that you’re comfortable with in the task checklist that your partner doesn’t. You speak in confidence when you tell your partner, oh no, you’ve repeated this word more than once; you’re not supposed to repeat a word more than once. Look at the task checklist; it says, you know, avoid repetition, so yeah, change these words because they’re the same. So, it also establishes; it gives them confidence, you know. So, these are the benefits that, in general, peer review produces. But because we’re marking this using a checklist and peer review, it’s a classic representation of what peer review is. And it’s fulfilling these benefits and objectives that they need to achieve.”

Through the observation, it was also found that students were able to give feedback to their peers using the checklist as the criteria of assessment. Students’ feedback centered on whether their peers’ performance was based on the specific criteria desired rather than on whether they had reached a certain grade or mark.

Following the assessment, the teacher asked the students how the assessment had helped them. One of the students answered that the assessment had helped them understand how to work as a group and cooperate. Another student responded that it had improved their ability to work as a team. These responses indicated that students felt they had become better at working with one another as a result of the task.

7.12. Utilization of Learners’ High-Order Thinking

Before writing about the issue, students had to determine the causes and effects of glaciers melting. They were able to accurately use the information they collected from the text and transfer it into writing on their own. In the second task, students were asked to respond to an article, which also prompted them to use their critical thinking to respond to the task. Students were also able to be critical of their performance and their peers’ performance.

7.13. Active Learning

During the observation, it was apparent that assessment prompted students to look for what they were miss-

ing during the task, as students who did not understand some aspects of writing (such as a topic sentence or a concluding sentence) began asking their peers and, at times, the teacher. Furthermore, throughout the reading portion of the task, students were able to use different types of strategies, such as underlining, note-taking, Internet searches, and peer consultation.

Following the assessment, the teacher asked students: What did you learn? The learners' responses included experiencing vocabulary expansion, understanding paragraph structure, improving teamwork skills, learning to work under time constraints, and realizing the value of peer assessment.

7.14. Improvement in Students' Writing

Some of the 10 students showed slight improvement from the first to the second task on aspects of the criteria.

Some students showed improvement in the areas of task fulfillment, organization, and coherence.

In all the figures, the sheet on the left shows students' response to first task, while the sheet on the right is students' response to second task.

Student A showed slight improvement in the areas of organization and coherence. The student's use of supporting sentences in the first task lacked depth and clarity, as he had failed to elaborate. In contrast, the student on the second task used supporting sentences more effectively, providing more specific statements and using supporting details (specifically, when he used the example from the Saudi culture to strengthen his argument). Additionally, the student showed specific improvement in spelling the word "what," as he had spelled it "wate" on the first task, which indicated that he avoided this mistake on his second task by spelling it correctly (see **Figure2**).

Glaciers melting is a huge
 Problem in our planet. so wate is the
 reason for a glaciers melting. one of the
 reasons is climat change. when the
 climat chang the heat in the earth change
 too. In short, I show you one reason
 for the a glaciers ~~melting~~
 melting.

The main reason for glaciers melting
 is human. we use fuel alot and we produce
 carbon dioxide. And thats effect in glaciers
 melting. And thats make the climat change
 changing more and more. In short, we effecting in
 ouer earth whith fuel thats the main reason for
 glaciers melting.

Write a response paragraph to the article above addressing the main points of the article.

I totally agree in what the ^{writer} says about
 wearing ~~an~~ traditional cloths. Many communities
 has their own traditional cloths and that cloths
 reflect for what they are. They wear their
 traditional cloths in a important events like
 wedding. for instans in sudi in the wedding,
 the guy how married should wear something called
 Bshit. And thats what I see.

Figure 2. Student A's Writing Performance on Tasks One and Two.

Student B showed improvement on task fulfillment (see **Figure3**). On the first task, he ignored the non-human causes of glacier melting and its effects on wildlife. However, on the second task, he fulfilled the task by responding to the main point of the article. The student also slightly improved in the use of supporting sentences and

details. In the first task, the supporting sentences were somewhat unclear, but on the second task, the student used more detailed and specific supporting sentences to discuss the points mentioned in the article. As is apparent on the second task, the student provided supporting details that were more coherent and relevant than those presented in

the first task.

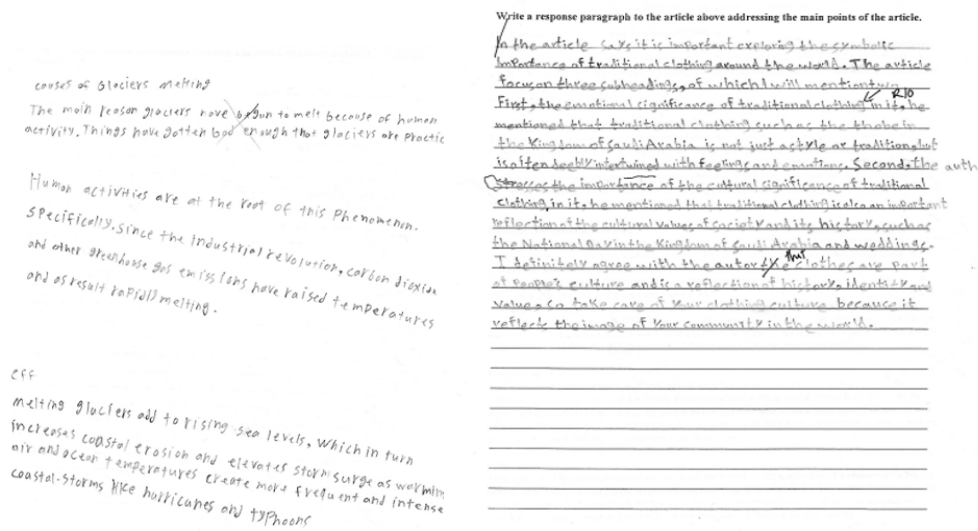


Figure 3. Student B's Writing Performance on Tasks One and Two.

In the first task, student C failed to address the non-human causes of glacier melting and its effects on wildlife (see Figure 4). In addition, he did not separate his response into two paragraphs, as the task required. However, on the second task, he responded to the main

point of the article about the symbolic significance of traditional clothing in one paragraph, as the task required. The student also effectively connected the discussion of traditional clothing, its societal significance, and personal agreement with the topic.

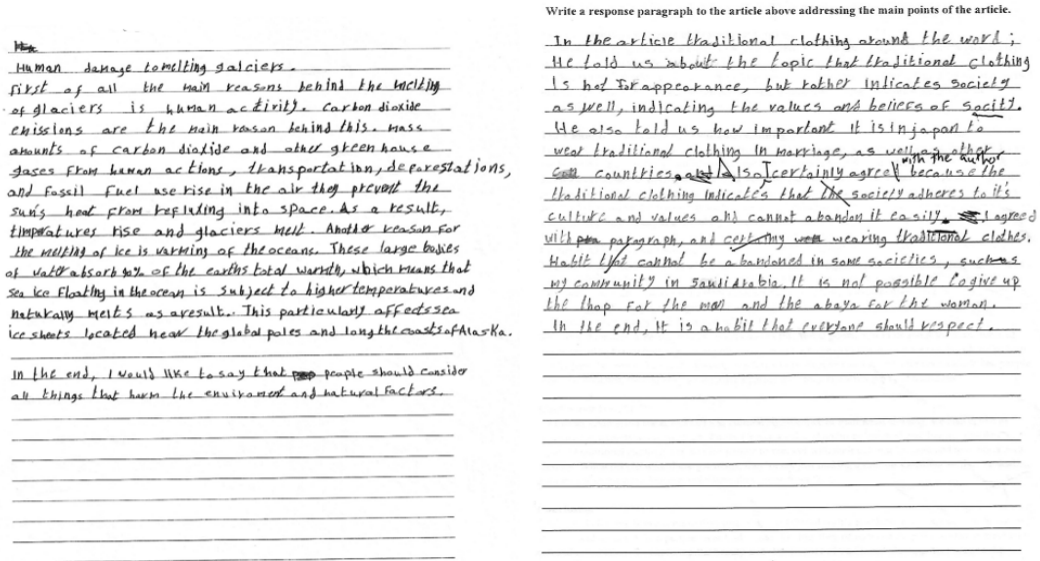


Figure 4. Student C's Writing Performance on Tasks One and Two.

Student D performed poorly on the first task (see Figure 5), but his task fulfillment improved noticeably. While his initial response was a scrambled version of the

text provided with Task One, it lacked clarity. The student's response on Task Two, however, fulfilled the task requirements and demonstrated that the student had put in

more effort this time. The student also showed considerable improvement in coherence and organization. Unlike

his performance on Task One, he maintained a logical and coherent flow of ideas on the second task.

to begin have glaciers reason
mainly today rising the activity
human of because is melt reason
primary the Earth the of temperature
this and more melt to started have
glaciers to back the directly because
change climate enough bad gotten have
things.

Studies show that only 2% of the
water available is fresh water
that people can consume over 70%
consists of glaciers and snow.

Write a response paragraph to the article above addressing the main points of the article.

My opinion about traditional clothes
is that they are significant and
should be respected. Traditional
clothing serves as a heritage that
has been passed down through generations.
Traditional clothing is a symbol of
pride and unity in many cultures.
In conclusion, the significance of
traditional clothes should not be
underestimated or disregarded.
Traditional clothing is an essential
aspect of cultural heritage and should
be respected and appreciated for its
unique beauty and value.

Figure 5. Student D's Writing Performance on Tasks One and Two.

As detailed in Figure 6, Student H showed improvement when it came to paragraph structuring from the first to the second task, as he presented his writing in an organized paragraph. In the second task, the learner provided a

response regarding the fabric of traditional clothes. However, the article was mainly concerned about the emotional and symbolic importance of traditional garments.

CAUSES of melting glaciers:

First, we have many reasons to melt of glaciers.
Case in point, earth's temperature is rising due to the
large number of factories. After that, factories produce
fumes and cause global warming.

Moreover, with global warming, there is damage to the
icy regions. As a result, the glaciers are melting.

On the other hand, there are many efforts by experts
to reduce global warming. In my opinion, I think
one of the most important reasons for reducing solubility
is to accompany the environment.

Effects of melting ice on humanity and the environment:

Undoubtedly, there will be environmental effects.
The poles will melt, and when this happens, the proportion
of water in the earth will increase.

On the other hand, many animals that live in ice will
become extinct. Besides, fresh water will disappear at a
high rate.

In conclusion, we must know that we are exposed to these
disasters because of global warming, but we don't in the
future to reduce the rate of warming.

Write a response paragraph to the article above addressing the main points of the article.

First, we all agree that there is no importance on the
type of fabric you wear. In addition to that, at this time,
the type of fabric doesn't matter, but rather the appearance
and consistency of the garment according to the place, the weather,
and the person. On the other hand, I agree in the opinion of the
writer to wear the traditional clothes of the city as for most
people, for example the KSA wearing a thobe is a traditional
garment that we wear on occasions such as marriage, holidays,
weddings or invitations. To summarize, many people don't care
about the type of fabric, but some of them are interested because
of the nature of his work or to sell it to people who care.

Figure 6. Student H's Writing Performance on Tasks One and Two.

Although Student J did not show improvement per se, he maintained excellent writing on both tasks (see **Figure 7**).

No noticeable improvement occurred. He completed both writing tasks in an organized and coherent manner.

The Arctic could be completely vacant of ice in less than 20 years, and that is due to melting ice glaciers which are melting at a rapid rate causing the whole world to be affected in the terrible situation. The main reason this is all happening has come down to human activity primarily. These activities include many things that we humans do as a daily task such as transportation whether it's using a car to travel from point a to b, or traveling with a plane, and another activity is fossil fuel usage which it also leads to transportation which all in the end lead to climate change. As a result, carbon emissions have become the main culprit of climate change which is itself being caused ultimately by humans. If humans do not try to reduce these emissions by any way possible we may see the Arctic naked of all its ice in the very near future.

Ultimately, the melting of all these ice glaciers has led to numerous consequences. Firstly, countless animals have suffered as a result of ice glaciers melting. Animals such as fish and polar bears have been getting the heat. The reason that water temperatures and levels have risen drastically, which in turn causes the survival rate of these animals. Furthermore, another major melting ice glaciers affects the future of the environment as a whole. It leads to water. Both water availability is limited, since most water is currently made up of ice glaciers and snow, this can be very problematic. Consequently, the future in regards of freshwater availability may very well be doomed if humans don't act on climate change.

Write a response paragraph to the article above addressing the main points of the article.

The importance of traditional clothing is made clear in the article. It is not just important culturally, but evolutionarily; it is a way for people to express their identity and connect with one another's cultural heritage. I wholeheartedly agree with the writer's view on traditional clothing. Firstly, traditional clothing plays a huge role evolutionarily in the individuals that wear them; hence, making them feel a sense of location and linkage to their cultures and traditions. Additionally, traditional clothing reflects the wearers' values and cultural significance. Moreover, the article also points out various traditional clothes are not just a piece of fabric, but a symbol of culture and history. These clothes are passed down from our ancestors or ancestors, making the value and importance of traditional clothing much more meaningful and valuable. Finally, I believe any form of traditional clothing is anything but just a mere piece of fabric; it is a reflection of history, culture, identity, and values.

Figure 7. Student J's Writing Performance on Tasks One and Two.

8. Discussion

The obtained results indicate that TBLA is thoroughly aligned with the syllabus objectives. The teacher's interview responses clearly hint at that. To put this into perspective, it is helpful to reference the design of TBLA. In the present study, TBLA was intended to be a formative assessment and an assessment for learning. A formative assessment involves the process of evaluating learners while they are learning to assist them in developing their abilities [5,32]. Thus, alignment with the syllabus is an important factor in fulfilling its function. In the present study, TBLA was also criterion-referenced. The criteria against which students' performance was measured were derived from the goals and objectives of the syllabus. Students' performance was assessed using a checklist that contained the criteria their performance should meet. An additional, the notable factor is TBLA itself and its attributes that contributed to the alignment with the objectives of the syllabus. Thus, it is clearly in alignment with the syllabus that endorses language use. This is in agreement with Norris [12,33] who stated that TBLA aligns with curricula that focus

on language use.

TBLA helped the teacher capture learners' actual writing performance because he was unsure whether some learners were actually doing their online blackboard assignments. This allowed him to provide personalized feedback to the learners. Using the information TBLA provided, the teacher could understand where learners could benefit from guidance and support, moving from targeting what learners were able to do on their own to the learners' ZPD, which is where learning occurs. This finding is significant because Vygotsky [34] defined the ZPD as "the distance between the actual development level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers" (p. 86).

The information received through TBLA prompted the teacher to act on his teaching practices to effectively align with the learners—moving from what learners can do on their own to their ZPD, as the teacher focused on shifting to what learners need guidance with to progress to the next stage in their learning. This finding was in line

with Liu ^[20], who implemented TBLA as a formative assessment in their context. They found that TBLA helped the teacher understand students' strengths and weaknesses in writing, allowing him to adjust his teaching. Similarly, Ke ^[17] reported that their criterion-referenced TBLA allowed their teachers to adjust their teaching and provide feedback to learners.

In social constructivism, learning occurs through social interactions ^[35]. The findings in this study showed that TBLA created a social environment where social interactions between learners and their peers were heightened. This created a space in which learners were the ones functioning as knowledgeable others to their peers. According to Vygotsky ^[34], social interactions with a more knowledgeable other are important in fostering learning. While learners were interacting with one another, they were guiding each other in their ZPD by scaffolding through feedback. These findings align with Brindley's ^[10] assertion that TBLA allows learners to acquire beneficial feedback that diagnoses their progress and accomplishments. This also aligns with Liu's ^[20] findings that the teacher was able to identify learners' strengths and weaknesses based on their performance on TBLA.

TBLA successfully went beyond learners' memorization and basic understanding to more complex cognitive processes, such as analysis, synthesis, and critical thinking. As learners engaged in authentic tasks, they were able to analyze, synthesize, and evaluate information, all of which processes are high-order thinking skills ^[36]. This lies in opposition to traditional assessment, which primarily requires lower-level thinking skills and more memorization ^[37]. TBLA is concerned with how learners apply knowledge to achieve tasks that resemble what is found in the real world ^[11]. The obtained results demonstrate that learners were able to use their language authentically by utilizing their linguistic and non-linguistic resources. During the implementation of TBLA, students showed increased motivation and engagement. Intrinsic motivation is defined by Ryan and Deci ^[38] as "doing of an activity for its inherent satisfactions rather than for some separable consequence" (p. 56). The findings indicate that TBLA helped create an environment facilitating intrinsic motivation. By introducing TBLA, learners were given autonomy

over their writing process.

This also indicates that learners were not externally pressured to produce accurate writing and avoid mistakes. This results from the fact that TBLA allows taking learning as a priority rather than accountability ^[39]. In our opinion, this removed the psychological pressure of attempting to get a high mark and allowed learners to feel at ease. Examining the results through the lens of social constructivism, learners functioned in their ZPD; those who showed such motivation and engagement transferred it to their classmates helping less motivated learners to learn how to become motivated. The assessment process triggered active learning. In the first TBLA implementation, students were given an integrated task that required them to read a text and write two paragraphs on the subject matter. Students did the reading in groups, and during this group activity, they were observed interacting with one another and helping each other understand the reading text before separately engaging in the writing. TBLA took students to their ZPD by challenging their current development. Moreover, although prior to each implementation the students were briefed about the task criteria and the assessment checklist and were asked if they had any questions, some still opted to ask these questions during the task. This means those learners were moving away from being passive learners to active learners.

While students engaged in the peer assessment following the writing stage, they were functioning in their ZPD. As each of them began to peer-assess, they exchanged feedback with each other, explaining to their classmates why their performance did not meet certain criteria. This demonstrated that students were scaffolding their peers to progress in their learning by diagnosing the areas that needed development using the checklist. This was in line with the essence of an assessment for learning, which is an assessment framework that prioritizes the facilitation of learning ^[5]. The finding also indicated that TBLA positively affected some students' writing on the second task. Those students showed a slight improvement in different aspects of their writing, indicating that the feedback provided by learners and the teacher-facilitated growth in those learners' ZPD was effective.

The improvement could also be a result of the wash-

back of the assessment process. Learners demonstrated slight improvements in meeting the task requirements and providing supporting details, both of which were criteria outlined in the checklist. The feedback provided to learners further suggested that TBLA played a role in facilitating this improvement, or at the very least, it contributed to it. The improvement some learners showed was a clear indication that there was a positive shift in their ZPD. According to learners' responses, they acquired new vocabulary, improved their spelling skills, gained an understanding of a topic sentence, and learned how to structure a paragraph. So, TBLA was not only an assessment evaluating learners' writing skills, but it also helped facilitate learning those skills, acting as an assessment for learning.

9. Conclusions

To sum up, the results obtained from data analysis can be effectively put forward to answer the research questions. The first research question was: How Effective is TBLA in Measuring Students' Writing Skills?

The obtained results show that TBLA is very effective in measuring students' writing skills. The teacher could capture learners' actual writing performance. Based on that, the teacher provided personalized feedback to his students. The information received through TBLA helped the teacher understand where guidance and support would benefit the learners, and what learners were able to do on their own. This finding is very significant. TBLA successfully went beyond learners' memorization and basic understanding to more complex cognitive processes, such as analysis, synthesis, and critical thinking used in writing. One of the benefits of TBLA observed in the results is the alignment with the goals and objectives of the course, which gives the assessment validity.

The second research question was: How Does TBLA Affect Students' Learning of Writing? The answer to that question is that TBLA encourages teamwork, exposes the students' actual performance in writing, reinforces the learning objectives in writing, help teachers provide personalized feedback to their students, has positive wash-back effect on students' learning, helps students identify their learning gaps in writing, provides better motivation

and engagement, being a student-centered approach it give more freedom to students to express themselves in writing, encourages teamwork and helps learners become high-order thinkers to become better writers.

10. Pedagogical Implications

The findings of the present study have pedagogical implications. The obtained results demonstrate that in addition to sufficient time, TBLA requires a classroom with movable chairs. The implementation of TBLA can exceed two hours depending on the task requirements. Language teachers should consider these issues before implementing the assessment. Additionally, TBLA showed high alignment with syllabus objectives. However, this alignment underscores that TBLA can be implemented only into a syllabus that reflects such a perspective. TBLA elevates the quality and value of the teacher's feedback. TBLA improves the feedback quality giving the teacher a realistic picture of learners' writing.

It should be emphasized at this juncture that assessment is a highly contextual process, and the needs of one context differ from those of another. When considering the findings of this study, educators should bear in mind that just because a benefit is found in one context does not necessarily imply that it will occur in their contexts as well, to avoid misalignment and undesirable consequences.

11. Limitations of the Study

Although the researcher made serious efforts, he could not achieve certain objectives for lack of time and resources, which may be considered as limitations of the study. First, the document analysis of students' writing was limited to only 10 students and was limited to the criteria predetermined by the researcher. Students' writing samples are largely representative, therefore, the small number is not a hindering issue. Second, considering the small number of participants, the findings of the present study cannot be generalized. However, since the research method is replicable, further research studies can explore the issue and contribute to generalizability. Third, all the participants in this study were male students. Fourth, the

obtained results were limited to two task implementations and did not illustrate the implications if more tasks were to be included. The findings also do not indicate whether the use of TBLA as a formative assessment influenced students' performance in summative tests. Lastly, the participants in the study represented only CEFR B1-level learners.

12. Recommendations for Future Research

It is recommended that future researchers consider longitudinal studies in the Saudi context in examining the impact of TBLA as a formative assessment. Future research studies can also focus on evaluating TBLA as a formative assessment of the entire language institution rather than just the classroom. Moreover, because assessment is highly contextual, researchers can also investigate whether implementing TBLA with students of other educational and language levels produces different results. Finally, researchers can examine TBLA's effectiveness as a formative assessment tool with different curriculum types to examine its suitability.

Author Contributions

Conceptualization, A.A. and M.A.; methodology, A.A.; formal analysis, A.A.; investigation, A.A.; resources, M.A.; data curation, M.A.; writing—original draft preparation, A.A.; writing—review and editing, M.A.; supervision, A.A. All authors have read and agreed to the published version of the manuscript.

Funding

This work received no external funding.

Institutional Review Board Statement

The study was conducted in accordance with the Declaration of Helsinki, and approved by the Ethics Committee of King Abdulaziz University (Ethics Committee 21 and date of approval October 20, 2022) for studies involving humans.

Informed Consent Statement

Informed consent was obtained from all subjects involved in the study.

Data Availability Statement

Data can be shared on request only with the permission of the Deanship of Scientific Research, King Abdulaziz University.

Conflicts of Interest

The authors declare no conflict of interest.

References

- [1] Black, P., Wiliam, D., 1998. Assessment and Classroom Learning. *Assessment in Education: Principles, Policies, and Practices*. 5(1), 7-74. DOI: <https://doi.org/10.1080/0969595980050102>
- [2] Coombe, C., Davidson, P., O'Sullivan, B., Stoyhoff, S. (eds.), 2012. *The Cambridge Guide to Second Language Assessment*. Cambridge University Press.
- [3] Noroozi, M., Taheri, S., 2022. Task-Based Language Assessment: A Compatible Approach to Assess the Efficacy of Task-Based Language Teaching vs. Present, Practice, Produce. *Cogent Education*. 9(1), 2105775. DOI: <https://doi.org/10.1080/2331186X.2022.2105775>
- [4] Alahmadi, N., Alrahaili, M., Alshraideh, D., 2019. The Impact of Formative Assessment in Speaking Tests on Saudi Students' Performance. *Arab World English Journal*. 10(1), 259-270. DOI: <https://doi.org/10.24093/awej/vol10no1.22>
- [5] Wiliam D., 2011. *Embedded Formative Assessment*. Solution Tree Press. <https://archive.org/details/embeddedformativ0000wili>
- [6] Sudharshana, N.P., Mukhopadhyay, L., 2021. *Task-Based Language Teaching and Assessment: Contemporary Reflections from Across the World*. Springer Singapore. DOI: <https://doi.org/10.1007/978-981-16-4226-5>
- [7] Van Gorp, K., Deygers, B., 2013. Task-Based Language Assessment. In: Kunnan, A.J. (ed.). *The Companion to Language Assessment: Vol. 2. Approaches and Development*. Wiley-Blackwell. 578–593. DOI: <https://doi.org/10.1002/9781118411360.wbcla058>
- [8] Brindley, G., 1994. Task-Centred Assessment in Language Learning Programs: The Promise and the Challenge. In: Bird, N., Falvey, P., Tsui, A., Allison,

- D., McNeill, A. (eds.). *Language and Learning*. Institute of Language in Education. 73–94. <https://eric.ed.gov/?id=ED386045>
- [9] Long, M.H., Norris, J.M., 2000. Task-Based Language Teaching and Assessment. In: Byram, M. (ed.). *Encyclopedia of Language Teaching*. Routledge. 597–603.
- [10] Brindley, G., 2012. Task-Based Assessment. In: Chapelle, C. (ed.). *The Encyclopedia of Applied Linguistics*. John Wiley & Sons. 1-6. DOI: <https://doi.org/10.1002/9781405198431.wbeal1141>
- [11] Ellis, R., 2003. *Task-based Language Learning and Teaching*. Oxford University Press.
- [12] Norris, J.M., 2016. Current Uses for Task-Based Language Assessment. *Annual Review of Applied Linguistics*. 36, 230–244. DOI: <https://doi.org/10.1017/S0267190516000027>
- [13] Bachman, L.F., 2002. Some Reflections on Task-Based Language Performance Assessment. *Language Testing*. 19(4), 453-476. DOI: <https://doi.org/10.1191/0265532202lt240oa>
- [14] Chalhoub-Deville, M., 2001. Task-Based Assessments: Characteristics and Validity Evidence. In: Bygate, M., Skehan, P., Swain, M. (eds.). *Researching Pedagogic Tasks: Second Language Learning, Teaching and Testing*. Routledge. 210–28.
- [15] Elder, C., Iwashita, N., McNamara, T., 2002. Estimating the Difficulty of Oral Proficiency Tasks: What does the Test-Taker have to Offer? *Language Testing*. 19(4), 347-368. DOI: <https://doi.org/10.1191/0265532202lt235oa>
- [16] Byrnes, H., 2002. The Role of Task and Task-Based Assessment in a Content-Oriented Collegiate Foreign Language Curriculum. *Language Testing*. 19(4), 419-437. DOI: <https://doi.org/10.1191/0265532202lt238oa>
- [17] Ke, C., 2006. A Model of Formative Task-Based Language Assessment for Chinese as a Foreign Language. *Language Assessment Quarterly*. 3(2), 207-227. DOI: https://doi.org/10.1207/s15434311laq0302_6
- [18] Sarıgöz, I.H., Fişne, F.N., 2019. Integration of Task-Based Language Assessment into Young Learner Classrooms. *Eurasian Journal of Applied Linguistics*. 5(3), 403-422. DOI: <https://doi.org/10.32601/ejal.651391>
- [19] Norris, J.M., 2009. Task-Based Teaching and Testing. In: Long, M.H., Doughty, C.J. (eds.). *The Handbook of Language Teaching*. Wiley-Blackwell. DOI: <https://doi.org/10.1002/9781444315783.ch30>
- [20] Liu, Q., 2018. Application of Task-Based Language Assessment in College English Writing Teaching. *Proceedings of the 2018 2nd International Conference on Economic Development and Education Management (ICEDEM 2018)*. 404–408. DOI: <https://doi.org/10.2991/icedem-18.2018.101>
- [21] Chen, S., Wang, J., 2019. Effects of Task-Based Language Teaching (TBLT) Approach and Language Assessment on Students' Competences in Intensive Reading Course. *English Language Teaching*. 12(3), 119-138. DOI: <https://doi.org/10.5539/elt.v12n3p119>
- [22] Almossa, S.Y., 2021. Assessment Practices in English Language Courses in Saudi Public Universities. *Collated Papers for the ALTE 7th International Conference*. Madrid, Spain. 175–178. <https://www.alte.org/IntConfProceedings>
- [23] Creswell, J.W., 2013. *Qualitative Inquiry and Research Design: Choosing among Five Approaches*, 3rd ed. SAGE.
- [24] Pathak, V., Jena, B., Kalra, S., 2013. Qualitative Research. *Perspectives in Clinical Research*. 4(3), 192. DOI: <https://doi.org/10.4103/2229-3485.115389>
- [25] Creswell, J.W., Poth, C.N., 2018. *Qualitative Inquiry and Research Design: Choosing among Five Approaches*, 4th ed. SAGE.
- [26] Dwyer, S.C., Buckle, J.L., 2009. The Space between: On Being an Insider-Outsider in Qualitative Research. *International Journal of Qualitative Methods*. 8(1), 54–63. DOI: <https://doi.org/10.1177/160940690900800105>
- [27] Brown, J.D., Hudson, T., 2002. *Criterion-Referenced Language Testing*. Cambridge University Press. DOI: <https://doi.org/10.1017/cbo9781139524803>
- [28] Creswell, J.W., Miller, D.L., 2000. Determining Validity in Qualitative Inquiry. *Theory into Practice*. 39(3), 124–130. DOI: https://doi.org/10.1207/s15430421tip3903_2
- [29] Gross J., 2018. Document Analysis. In: Frey, B. (ed.). *The SAGE Encyclopedia of Educational Research, Measurement and Evaluation*. SAGE. DOI: <https://doi.org/10.4135/9781506326139.n209>
- [30] Bowen, G.A., 2009. Document Analysis as a Qualitative Research Method. *Qualitative Research Journal*. 9(2), 27–40. DOI: <https://doi.org/10.3316/QRJ0902027>
- [31] Braun V., Clarke, V., 2006. Using Thematic Analysis in Psychology. *Qualitative Research in Psychology*. 3(2), 77–101. DOI: <https://doi.org/10.1191/1478088706qp063oa>
- [32] Brown, H.D., 2004. *Language Assessment: Principles and Classroom Practices*. Pearson Education.
- [33] Norris, J.M., 2018. Task-Based Language Assessment Aligning Designs with Intended Uses and Consequences. *JLTA Journal*. 21, 3-20. DOI: https://doi.org/10.20622/jltajournal.21.0_3
- [34] Vygotsky, L.S., 1978. *Mind in Society: The Develop-*

- ment of Higher Psychological Processes. Cambridge, MA: Harvard University Press.
- [35] Amineh, R.J., Asl, H.D., 2015. Review of Constructivism and Social Constructivism. *Journal of Social Sciences, Literature and Languages*. 1(1), 9–16. [https://www.blueap.com/j/List/4/iss/volume%2001%20\(2015\)/issue%2001/2.pdf](https://www.blueap.com/j/List/4/iss/volume%2001%20(2015)/issue%2001/2.pdf)
- [36] Miri, B., David, B.C., Uri, Z., 2007. Purposely Teaching for the Promotion of Higher-Order Thinking Skills: A Case of Critical Thinking. *Research in Science Education*. 37(4), 353–369. DOI: <https://doi.org/10.1007/s11165-006-9029-2>
- [37] Dikli, S., 2003. Assessment at a Distance: Traditional vs. Alternative Assessments. *The Turkish Online Journal of Educational Technology*. 2(3), 13–19. <http://www.tojet.net/articles/v2i3/232.pdf>
- [38] Ryan, R.M., Deci, E.L., 2000. Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being. *American Psychologist*. 55(1), 68–78. DOI: <https://doi.org/10.1037/0003-066X.55.1.68>
- [39] Black, P., Harrison, C., Lee, C., et al., 2004. Working Inside the Black Box: Assessment for Learning in the Classroom. *Phi Delta Kappan*. 86(1), 8–21. DOI: <https://doi.org/10.1177/003172170408600105>