

Assessment in Task-Based Language Teaching (TBLT): Challenges and Approaches to Overcome Them

Nargiza Djurayeva Kudratillayevna

“International School of Finance, Technology and Science” Institute

Abstract:

Assessment plays a critical role in Task-Based Language Teaching (TBLT) by determining the effectiveness of task execution and evaluating students' progress. However, assessing students in TBLT presents unique challenges, particularly when dealing with diverse groups with varying levels of English proficiency and social backgrounds. This article examines these challenges and explores approaches to address them, using an experiment conducted with banking and finance students at the International School of Finance Technology and Science in Tashkent as a case study. The findings highlight practical solutions to enhance the fairness and accuracy of assessment in TBLT. The study is supported by literature on assessment in language learning and pedagogical frameworks.

Keywords: assessment in TBLT, diverse classrooms, proficiency levels, formative assessment, summative assessment, differentiated rubrics, peer feedback, collaborative learning, equitable assessment practices.

Introduction

Task-Based Language Teaching (TBLT) has emerged as an effective methodology for language instruction, emphasizing the use of language in authentic and meaningful contexts. By focusing on real-world tasks, TBLT aligns with communicative language teaching principles and enables students to develop linguistic, cognitive, and practical skills simultaneously (Ellis, 2003). The methodology's student-centered approach encourages active participation and provides opportunities for students to apply language skills in realistic scenarios. However, one of the most pressing issues in TBLT is assessment. Effective assessment is essential for understanding student

progress, providing feedback, and ensuring the success of the teaching methodology. Yet, the dynamic nature of tasks and the focus on communication over linguistic accuracy pose significant challenges to creating fair and comprehensive assessments.

Assessing students within a TBLT framework becomes even more complex in diverse classrooms. Variations in English proficiency, cultural and social backgrounds, and learning styles can exacerbate these difficulties. Teachers must balance multiple objectives, including evaluating linguistic competence, task performance, and collaboration, while ensuring fairness and inclusivity (Skehan, 1998). In such settings, assessment frameworks must be flexible enough to accommodate diverse learners without compromising the integrity of the evaluation process.

This study examines the challenges associated with assessing students in TBLT by analyzing an experiment conducted with banking and finance students at the International School of Finance Technology and Science in Tashkent. The students in this group exhibited diverse English proficiency levels and came from varied social and economic backgrounds, making assessment particularly challenging. The study explores strategies to refine assessment practices in TBLT and aims to answer the following research questions: What are the primary challenges of assessing students in TBLT? How can assessment approaches in TBLT be refined to accommodate diverse student groups? What practical solutions can teachers implement to ensure fair and comprehensive evaluation?

Methodology

Research Design

The research adopted a qualitative case study approach to investigate the challenges of assessment in TBLT and identify practical solutions. The study was conducted over three weeks and included both formative and summative assessment components. Qualitative methods were employed to capture the nuances of assessment challenges and the effectiveness of implemented solutions. Data were collected through classroom observations, analysis of task outputs, teacher journals, and focus group discussions with students.

Participants

The participants comprised 20 undergraduate students specializing in banking and finance at the International School of Finance Technology and Science in Tashkent. The group represented a microcosm of diversity, with students exhibiting varying levels of English proficiency, ranging from beginner to advanced. Additionally, their social and economic backgrounds differed significantly, affecting their access to resources and prior exposure to English. This diversity provided a valuable context for examining the practical challenges of assessment in a TBLT framework.

Task Design

The task designed for this study required students to create a financial portfolio for a fictional client, simulating a real-world financial advisory scenario. The task was intentionally complex, involving multiple stages: researching market trends, analyzing client needs, drafting a financial report, and delivering an oral presentation. The design aimed to integrate subject-specific knowledge with language skills, ensuring that students engaged in authentic, contextually relevant activities.

The research phase required students to use resources provided by the teacher, such as case studies and market data. In the report-writing phase, students employed language tools like Grammarly to refine their writing, focusing on clarity and accuracy. During the presentation phase, students were encouraged to use visual aids and professional presentation tools, such as Canva or PowerPoint.

These steps provided opportunities for formative feedback while aligning with the communicative goals of TBLT.

Assessment Framework

The assessment framework in this study was designed to address the multidimensional nature of TBLT while accommodating the diverse needs of the participants. A combination of formative and summative assessment methods was employed to evaluate student performance comprehensively. Formative assessment took the form of ongoing teacher observations, peer feedback sessions, and mid-task evaluations, allowing the teacher to monitor progress and provide timely interventions. These formative components helped identify areas of difficulty and provided students with opportunities to improve before the final evaluation.

Summative assessment focused on the final task outputs, including written financial reports and oral presentations. The evaluation process relied on a detailed multi-criteria rubric, which was carefully developed to balance linguistic accuracy, task fulfillment, presentation skills, and evidence of collaboration. By incorporating specific descriptors for each criterion, the rubric ensured transparency and consistency in grading. For instance, the "Language Use" category evaluated accuracy, fluency, and appropriateness, while "Task Fulfillment" assessed the relevance, coherence, and depth of the financial portfolio.

Recognizing the diversity in students' proficiency levels, the rubric also featured tiered expectations. Beginner-level students were primarily assessed on their ability to convey basic ideas and demonstrate comprehension of the task, while advanced students were expected to exhibit critical thinking, professional articulation, and nuanced language use. This differentiated approach not only ensured fairness but also motivated students to perform at their optimal level within their respective abilities. Evidence from similar studies (Carless & Boud, 2018) supports the effectiveness of tiered rubrics in promoting equitable assessment outcomes.

Data Collection

The data collection process for this study was methodical and multi-faceted, designed to capture a holistic view of the assessment process and its challenges. Classroom observations formed a critical component, allowing the teacher to document real-time interactions, student behaviors, and emerging issues during task execution. Detailed observation notes provided qualitative insights into how students approached the task and how they responded to formative feedback.

Task outputs, including written reports and recorded presentations, were another vital data source. These artifacts were analyzed against the rubric criteria to identify patterns in student performance. For example, linguistic errors were more prevalent among beginner-level students, while advanced students showed greater variability in task-specific knowledge. Comparing outputs across proficiency levels enabled the teacher to evaluate the rubric's effectiveness and refine assessment strategies.

Focus group discussions provided an additional layer of depth by capturing students' perspectives on the assessment process. Conducted after the task was completed, these discussions encouraged students to share their experiences, challenges, and suggestions for improvement. The teacher used open-ended questions to prompt meaningful reflections, such as "How did the peer feedback process help you improve?" and "What aspects of the rubric did you find most helpful or challenging?" Transcripts from these sessions revealed valuable insights into students' perceptions of fairness and clarity in assessment.

Finally, the teacher maintained a reflective journal throughout the study, documenting daily observations, emerging themes, and personal reflections on the assessment process. This journal provided a meta-analytic perspective, enabling the teacher to identify recurring patterns and

evaluate the effectiveness of implemented strategies over time. Triangulating data from these diverse sources enhanced the reliability and validity of the study's findings.

Results

Challenges in Assessment

The findings revealed several interconnected challenges in assessing students within the TBLT framework. One significant challenge was the disparity in English proficiency levels among students. Advanced students demonstrated greater fluency and linguistic accuracy, enabling them to excel in both written and oral tasks. In contrast, beginners struggled with basic language structures, which hindered their ability to express complex ideas. This disparity created difficulties in applying uniform assessment criteria.

Another challenge stemmed from students' varied social and economic backgrounds. Those with limited access to resources, such as stable internet connections or academic tools, faced additional barriers. For example, some students were unable to conduct thorough market research due to resource constraints, resulting in uneven task outputs. These disparities highlighted the importance of providing equitable support to all learners.

The subjectivity of peer feedback also emerged as a concern. While peer evaluations were intended to foster collaboration and reflection, inconsistencies in feedback quality affected its reliability. Some students provided overly critical feedback, which demotivated their peers, while others gave overly lenient evaluations that lacked constructive input. This inconsistency underscored the need for structured guidance in peer assessment.

Finally, the dual focus of TBLT on language use and task performance posed challenges in balancing assessment criteria. Students with strong content knowledge but weaker language skills were at a disadvantage, as their linguistic errors often detracted from the overall quality of their work. Conversely, students with strong language skills but limited subject knowledge struggled to meet task-specific requirements, raising questions about the relative weighting of these criteria.

Strategies to Overcome Challenges

To address these challenges, the study implemented several strategies that proved effective. One key approach was the use of differentiated rubrics, which set tailored expectations for students based on their proficiency levels. By adjusting the weighting of linguistic and task-specific criteria, the teacher ensured that all students were assessed fairly and comprehensively. For example, beginner-level students received additional credit for effort and comprehension, while advanced students were challenged to demonstrate critical thinking and professional articulation.

Blended feedback mechanisms were another successful strategy. Peer feedback was supplemented with teacher feedback to provide balanced and actionable input. The teacher conducted training sessions to help students understand how to give constructive feedback, emphasizing the importance of specific, evidence-based comments. This approach not only improved the quality of peer evaluations but also fostered a collaborative classroom environment.

Scaffolding and resource support were also critical in addressing disparities. Students with limited access to resources were provided with supplementary materials, such as printed market data and access to language tools like Quillbot. These supports helped level the playing field and enabled all students to produce higher-quality outputs. For example, Grammarly's real-time suggestions on grammar and style improved the written reports of beginner-level students, while advanced students benefited from its feedback on nuanced language use (Wilson & Cardwell, 2020).

Adjusting the weighting of assessment criteria was another practical solution. To ensure fairness, linguistic accuracy was given slightly less weight than task fulfillment and professional skills. This

adjustment acknowledged the dual goals of TBLT without penalizing students disproportionately for their linguistic limitations.

Student Perceptions

Student feedback revealed a range of perspectives on the assessment process, offering valuable insights into its perceived strengths and weaknesses. Many students expressed appreciation for the transparency of the rubric, which clarified expectations and provided a roadmap for task completion. Advanced students noted that the rubric's emphasis on critical thinking and professional skills encouraged them to push beyond basic task requirements and refine their outputs. Beginner-level students, on the other hand, valued the differentiated criteria, which acknowledged their efforts and enabled them to focus on manageable goals.

The inclusion of peer feedback was another aspect that received mixed reviews. Some students found the peer evaluation process empowering, as it allowed them to engage critically with their peers' work and reflect on their own. However, others highlighted issues with feedback consistency, noting that overly critical or overly lenient evaluations undermined the process's effectiveness. These observations echo findings by Carless and Boud (2018), who emphasize the need for clear guidelines and training to improve peer feedback reliability.

Students also appreciated the additional scaffolding provided to address resource disparities. Those with limited internet access or less familiarity with digital tools reported that the supplementary materials and support bridged critical gaps. For example, printed data sets and guided tutorials on tools like Canva enabled them to contribute meaningfully to the task despite initial disadvantages. Overall, the feedback indicated that while the assessment process was challenging, the implemented strategies helped create a more equitable and supportive learning environment.

Discussion

The challenges identified in this study align with existing literature on TBLT assessment. Skehan (1998) highlights the inherent tension between evaluating linguistic accuracy and task performance, which becomes particularly pronounced in diverse classrooms. The findings also support Ellis's (2003) argument that assessments in TBLT should prioritize communicative effectiveness and task completion over traditional language testing metrics.

Differentiated rubrics emerged as a particularly effective tool for addressing diversity in proficiency levels. By tailoring expectations to individual student needs, teachers can create a more inclusive and equitable assessment framework. This approach aligns with Carless and Boud's (2018) emphasis on the importance of flexible assessment practices that accommodate learner diversity.

The integration of blended feedback mechanisms also holds significant promise. Training students to provide constructive peer feedback not only enhances the quality of formative assessment but also fosters a sense of responsibility and collaboration. Zawacki-Richter et al. (2019) highlight the potential of such approaches to build students' feedback literacy and improve overall learning outcomes.

However, the study also underscores the need for further research into the long-term impact of these strategies. Longitudinal studies could explore how differentiated assessment practices influence student motivation, performance, and skill development over time. Additionally, future research could examine the applicability of these strategies across different educational contexts and disciplines.

Conclusion

Assessment in TBLT presents unique challenges, particularly in diverse classrooms where students vary in proficiency and background. The experiment with banking and finance students at the

International School of Finance Technology and Science highlights both the difficulties and potential solutions in this area. By employing differentiated rubrics, blended feedback, and integrated scaffolding, teachers can address the complexities of TBLT assessment while ensuring fairness and inclusivity. As the field continues to evolve, further research is needed to refine assessment practices and support the implementation of TBLT in diverse educational contexts.

References

1. Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. *Assessment & Evaluation in Higher Education*, 43(8), 1315-1325.
2. Ellis, R. (2003). *Task-based language learning and teaching*. Oxford University Press.
3. Skehan, P. (1998). *A cognitive approach to language learning*. Oxford University Press.
4. Wilson, S., & Cardwell, J. (2020). Improving writing proficiency through AI-driven feedback: A case study. *Language Learning & Technology*, 24(2), 12-25.
5. Willis, D., & Willis, J. (2007). *Doing task-based teaching*. Oxford University Press.
6. Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education: The changing landscape of teaching and learning. *International Journal of Educational Technology in Higher Education*, 16(1), 39.
7. Zhang, Y., & Elder, C. (2011). Judgments of oral proficiency by non-native and native English-speaking teacher raters: Competing or complementary constructs? *Language Testing*, 28(1), 31-50.