

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/341030611>

Task-Based Language Assessment: Components, Development, and Implementation

Chapter · April 2020

CITATIONS

14

READS

2,961

1 author:



Ali Shehadeh

United Arab Emirates University

50 PUBLICATIONS 702 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:



Currently I am working on a large joint international encyclopedia project on TESOL and on 2-3 papers on TBLT. [View project](#)

CHAPTER | 8

Task-Based Language Assessment: Components, Development, and Implementation

Ali Shehadeh

INTRODUCTION

The main purpose of this chapter is to explain testing / assessment within the framework of task-based language assessment (TBLA) and the wider context of task-based language teaching (TBLT). The chapter will describe TBLA and its characterizing features, illustrating the main components of TBLA, its development, and its classroom implementation and utilization. In order to properly understand and contextualize TBLA, however, we must first understand TBLT.

WHAT IS TASK-BASED LANGUAGE TEACHING, OR TBLT?

TBLT is an educational framework and an approach for the theory and practice of second / foreign language (L2) learning and teaching, and a teaching methodology in which classroom tasks constitute the main focus of instruction (Richards et al., 2003). A classroom task is defined as an activity that is (1) goal-oriented, (2) content-focused, (3) has a real outcome, and (4) reflects real-life language use and language need (Shehadeh, 2005). The syllabus in TBLT is organized around activities and tasks rather than in terms of grammar or vocabulary (Richards et al., 2003).

The interest in TBLT is based on the strong belief that it facilitates second language acquisition (SLA) and makes L2 learning and teaching more principled and more effective. This interest arose from a constellation of ideas issuing from philosophy of education, theories of SLA, empirical findings on effective instructional techniques, and the exigencies of language learning in contemporary society (Van den Branden et al., 2009). It is beyond the limits of this chapter to illustrate the various perspectives on TBLT and ways of implementing and utilizing it in the L2 classroom. (For reviews, see, for instance, Ellis, 2003; Shehadeh and Coombe, 2010; Van den Branden et al., 2009.)

Indeed, tasks are now viewed as “the devices that provide learners with the data they need for learning” (Ellis, 2000: 193). No wonder, therefore, that many teachers around the world are moving toward TBLT, and no wonder that task-based pair and group work is now considered a standard teaching / learning strategy in many language classrooms around the world (Shehadeh and Coombe, 2010).

After this brief introduction to TBLT, it is now possible to describe task-based language assessment, its defining characteristics, its basic components, the basic steps in its development and implementation, and its classroom application and utilization.

WHAT IS TASK-BASED LANGUAGE ASSESSMENT, OR TBLA?

TBLA is a framework for language testing / assessment that takes the task as the fundamental unit for assessment and testing. It is based on the same underlying principles as TBLT, but extends them from the learning-and-teaching domain to the testing domain. Specifically, as in TBLT methodology, testing / assessment in TBLA is also organized around tasks rather than in terms of grammar or vocabulary. For instance, Long and Norris (2000: 600) state that “genuinely task-based language assessment takes the task itself as the fundamental unit of analysis, motivating item selection, test instrument construction and the rating of task performance.” As in TBLT methodology too, the main goal and validity of TBLA is measured against the extent to which it can successfully achieve a close link between the testee’s performance during the test and his / her performance in the real world. For instance, Ellis (2003: 279) states that “task-based testing is seen as a way of achieving a close correlation between the test performance, i.e. what the testee does during the test, and the criterion of performance, i.e. what the testee has to do in the real world.” Assessment tasks are thus viewed as “devices for eliciting and evaluating communicative performances from learners in the context of language use that is meaning-focused and directed towards some specific goal” (ibid.).

Four main features characterize TBLA. *First*, it is a formative assessment; that is, it is an assessment undertaken as part of an instructional program for the purpose of improving learning and teaching. *Second*, it is a performance-referenced assessment; that is, it is an assessment that seeks to provide information about learners’ abilities to use the language in specific contexts, that is directed at assessing a particular performance of learners, and that seeks to ascertain whether learners can use the L2 to accomplish real target tasks. *Third*, it is a direct assessment; that is, it is an assessment that involves a measurement of language abilities that involves tasks where the measure of the testee’s performance is incorporated into the task itself, like information-transfer test tasks such as information-gap, opinion-gap, and reasoning-gap tasks. It must be noted, however, that direct assessment still involves some level of inferring because it is necessary to observe performance and then infer ability from that performance. Put differently, you can measure outcomes, but you are still left with inferring the ability that produced the outcome. *Fourth*, it is an authentic assessment; that is, it is an assessment that involves either real-world language use (or as close as possible to this), or the kinds of language processing found in real-world language use, that is, the test task’s characteristics must match those of the target-language task (Ellis, 2003: 285).

COMPONENTS OF TASK-BASED LANGUAGE ASSESSMENT

Task-based language assessments consist of three basic components: a test task, an implementation procedure, and a performance measure:

TEST TASK DESIGN AND SELECTION

It is possible to identify two approaches to test task design and selection within the framework of TBLA: The first is called the *construct-centered approach* or the direct system-referenced tests. This approach entails the identification of a theory of language learning and language use as a basis for the test task design and selection. This approach is used with tasks that seek “to establish the general nature of the testees’ language proficiency” (Ellis, 2003: 286). The second approach is called the *work-sample approach* or the direct performance-referenced tests. This approach involves analysis of the target situation to determine what tasks the testee will need to perform in the real world. This approach is used with tasks that seek “to find out what a learner can do in a particular situation” (Ellis, 2003: 286). It should be noted, however, that the two approaches must be seen as complementary rather than mutually exclusive because any successful TBLA ideally requires both approaches.

IMPLEMENTATION PROCEDURES

There are two implementation procedures: The first is *planning time*. Planning time must be adopted as a key implementation procedure because it can improve the test taker’s performance. The second procedure is the *interlocutor* (on oral test tasks). The role of the addressee (familiar or unfamiliar, native speaker or nonnative speaker) plays a significant role in the performance of the testee on an assessment task. For instance, some researchers have shown that if you want to elicit the “best performance” from the testee, it may be preferable to set up a testing situation where the candidates interact with another nonnative speaker than with a native speaker (e.g., Wigglesworth, 2001). Along the same lines, one wonders whether it is also possible to have real recipients for writing tasks, analogous to the interlocutor in oral tasks. This is an important question for further exploring the implementation of TBLA in the L2 classroom.

MEASURING PERFORMANCE

How do we assess the performance elicited from the testee in TBLA tests? Two principal methods for measuring learner performance are used in TBLA contexts. The first is the *direct assessment of task outcomes*. This method involves either the assessor observing a performance of a task and making a judgment, or no judgment on the part of the assessor (i.e., judgment / measurement is objective – the testee did or did not succeed in performing the task) because it results in solutions that are either right or wrong. A good example of direct assessment of task outcomes is a closed task that results in a solution that is either right or wrong, like locating a journal article relating to a particular topic in a library (Robinson and Ross, 1996: 474–475). The main advantages of this method are that “it affords an objective measurement, involving no judgment on the part of the assessor, and it is easy and quick” (Ellis, 2003: 296).

The second performance measure is *external ratings*. This method involves external judgment which is more subjective (but efforts are being made to make it more reliable). This measure also involves a holistic measure (scale) or an analytic measure (scale) of linguistic ability (e.g., paragraph or essay evaluation; overall linguistic ability versus the four language skills identified for rating separately). Examples of this assessment include judging that the learner’s speaking ability is at the “expert” level in the oral interview component of the old ELTS test which specifies that the learner “can speak with authority on a variety of topics; can initiate, expand, and develop a theme” (Ellis, 2003: 300). The main advantage of this method is that it enables the assessor to specify the learner’s

competencies to be measured in more functional terms. Competencies in external ratings are determined in terms of learner performance levels or checklists (for fuller discussion, see Ellis, 2003: 296 ff.)

Along the same lines, Long and Norris, bearing in mind language professionals and classroom teachers, specify six practical steps for developing and implementing task-based language tests as follows:

1. The *intended use(s)* for task-based assessment within the language programme must be specified, minimally addressing the following four issues: who uses information from the assessment? [. . .]; what information is the assessment supposed to provide? [. . .]; what are the purposes for the assessment? [. . .]; and who or what is affected, and what are the consequences of the assessment?
2. Target tasks or task-types emerging from the needs analysis are analyzed and classified according to a variety of *task features*. Analysis is undertaken in order to understand exactly what real-world conditions are associated with target tasks and should therefore be replicated under assessment conditions.
3. Based on information from the analysis of task features, *test and item specifications* are developed. . . Specifications delineate the formats tests should take, procedures involved, tasks or task-types to be sampled, format for test tasks (items), and how performance on the task-based test should be evaluated.
4. [Carrying out] identification and specification of *rating criteria*, which form the basis for interpretations of examinee performance and task accomplishment. . . Real-world criterial elements (aspects of task performance that will be evaluated) and levels (descriptions of what success looks like on these aspects of task performance) should be identified within initial needs analysis, with a view toward providing students and teachers with clear learning objectives.
5. Task items, test instruments and procedures and rating criteria need to be *evaluated* (involving pilot-testing and revision) according to their efficiency, appropriacy and effectiveness with respect to the intended assessment uses.
6. Finally, task-based language assessment should incorporate procedures for systematic and ongoing *validation* of its intended use within the language programme. Validation should minimally consider: to what extent test instruments and procedures are providing appropriate, trustworthy and useful information; to what extent particular uses for the assessment are warranted, . . . and to what extent the consequences of assessment use can be justified, given the impact on students, teachers, language programmes and any other relevant stakeholders in the assessment process. (Long and Norris, 2000: 600–601)

UTILIZATION AND CLASSROOM APPLICATIONS OF TASK-BASED LANGUAGE ASSESSMENT

TBLA has been utilized by language teachers in the L2 classroom in a broad range of formal and informal educational settings that serve a wide range of language learners who come from different age groups, different proficiency levels, and different educational and cultural backgrounds. For space limitations, I will only cite in this section

selective, but illustrative, examples that demonstrate recent applications of TBLA in the L2 classroom.

Gan, Davison, and Hamp-Lyons (2009) analyzed the discourse derived from a task, which in their particular case was based on choosing a gift for the main character in the movie *Forrest Gump*. They found that peer group discussion as an oral assessment format has the potential to provide opportunities for students to demonstrate “real-life” spoken interactional abilities. Yeh (2010) describes a blended task-based activity designed for intermediate- and advanced-level students in speech or oral training classes. She describes a number of assessment tools used to evaluate student performance. Winke (2010) explains how to use online tasks for formative language assessment, which is defined as assessment that is used in “evaluating students in the process of ‘forming’ their competencies and skills” (Brown, 2004: 6). She demonstrates how these tasks provide continuous feedback to the teacher and learners, and how this feedback can be used for making decisions about ongoing instructional procedures and classroom tasks.

In a new collection of studies on TBLT and TBLA (see Shehadeh and Coombe, forthcoming, 2012), Weaver (forthcoming), like Winke (2010) above, provides an account of incorporating a formative assessment cycle into a TBLT curriculum in relation to a task for Japanese business students with a strong sense of authenticity.

It is important to note that in all these contexts, and in line with the underlying principles of TBLT and TBLA, assessment has been utilized and implemented not as an isolated component, but rather as *part and parcel* of the learning-and-teaching process. That is, assessment was intended to both facilitate learning and provide feedback that can be used for making further decisions about instructional procedures and teaching methodology. More specifically, assessment has been undertaken as part of the instructional program for the purpose of improving learning and teaching. At the same time, such assessments are both authentic and involve a measurement of language abilities that involve tasks where the measure of the testee’s performance is incorporated into the task itself (see characteristics of TBLA earlier in the chapter).

CONCLUSIONS: FUTURE DIRECTIONS

Carless (forthcoming) argues that “the use of teacher assessment in a traditionally examination-oriented system is both a powerful impetus for change and a challenge to teachers’ workloads and mindsets.” It is a challenge because traditionally examination-oriented systems have a long history of public assessment. In such systems, the uniform test, which emphasizes grammar, vocabulary, and reading comprehension, can be an obstacle to the implementation of TBLA (Ng and Tang, 1997). Language tests in these systems play a key part in success in the pursuit of personal well-being and future employment opportunities. According to Cheng (2008), for instance, the high status of English tests in China does not support teaching but drives it. Carless points out that in such settings, “there is a likelihood that test takers (often with the assistance of the ubiquitous after-hours tutorial sector) may seek to subvert the aims of the test developers.” He cites an example from a recent study (Luk, 2010) which reveals that “a careful analysis of the discourse of test interaction . . . shows how students colluded in producing utterances aimed at creating the impression of being effective interlocutors for the purpose of scoring marks rather than for authentic communication” (Carless, forthcoming).

The use of teacher assessment in traditionally examination-oriented systems, on the other hand, can be an impetus for changing teachers’ mindsets because it is proving again and again its futility and thus falling into disfavor, instead favoring TBLA which

is a more powerful assessment tool, as described above. No wonder, therefore, that high-stakes examinations have over the last twenty years increased the weighting awarded to oral performance and the examinations have become increasingly task-based (Carless, forthcoming).

Nonetheless, more research still needs to be done into TBLA. On the research side, we still need to know, for instance, to what extent does a formative assessment cycle support the development of wider student learning capacities, or more narrowly help them to improve their performance (Carless, forthcoming). As Black and Wiliam (2003) have argued, an assessment sequence can only be said to have acted formatively if it advances student learning. It may be that a way forward for formative assessment in relation to TBLA is to engage more with the expanding educational literature on formative assessment (e.g., Andrade and Cizek, 2010) and developments in dynamic assessment (e.g., Poehner, 2008).

On the pedagogical side, TBLA – as well as its underlying principles – should be made more accessible to practicing teachers. One reason why traditional, examination-oriented systems still persist in many parts of the world is because many teachers do not know how to utilize TBLA in their practices. Another reason is that not many teachers or instructors, when pressed hard, know what exactly is meant by TBLA and why it is more conducive to L2 learning. Also, many teachers, particularly in EFL / ESL contexts, believe that tasks are another face of the traditional exercises. Finally, many teachers consider TBLA to be an “alien theoretical concept” that is not applicable to their specific teaching context or educational setting.

Obviously, more work and more research need to be done to combat these hindrances in the design, implementation, application, and utilization of TBLA. As Brindley stated more than fifteen years ago (Brindley, 1994: 90), the challenge for researchers and test designers in TBLA is to develop a method that provides valid and reliable measurements in assessments that are “complex, qualitative, and multidimensional, rather than uniform and standardized.” This challenge continues to manifest itself in one form or another. Nevertheless, it is worth doing in order to arrive at an accurate assessment of the test taker’s ability, which is a crucial aspect of task-based language assessment.

Suggested resources

- Brindley, G. (2008). Educational reform and language testing. In E. Shohamy & N. Hornberger (Eds.), *Encyclopedia of language and education*. Vol. 7: *Language testing and assessment* (pp. 365–378). New York: Springer.
- Brown, J. D., Hudson, T., Norris, J., & Bonk, W. (2002). *An investigation of second language task-based performance assessments*. Honolulu, Hawai'i: University of Hawai'i Press.
- Long, M. H., & Norris, J. M. (2000). Task-based language teaching and assessment. In M. Byram (Ed.), *Encyclopedia of language teaching* (pp. 597–603). London: Routledge.
- Norris, J. M. (Ed.) (2002). Special issue: Task-based language assessment. *Language Testing*, Vol. 19, Issue 4.
- Norris, J. M. (2009). Task-based teaching and testing. In M. Long & C. Doughty (Eds.), *Handbook of second and foreign language teaching*. Cambridge: Blackwell.
- Shehadeh, A., & Coombe, C. (Eds.) (2010). *Applications of task-based learning in TESOL*. Alexandria, VA: TESOL Publications.
- Van den Branden, K., Bygate, M., & Norris, J. M. (Eds.) (2009). *Task-based language teaching: A reader*. Amsterdam: John Benjamin's.

Wigglesworth, G. (2008). Tasks and performance-based assessment. In E. Shohamy & N. Hornberger (Eds.), *Encyclopedia of language and education: Language testing and assessment*. Vol. 7: *Language testing and assessment* (pp. 111–122) (2nd ed.). New York: Springer.

Discussion questions

1. Do you utilize TBLA in your specific educational context? Why? Why not? What are the challenges?
2. Does your institution encourage assessment procedures alternative to an examination-oriented system? Why? Why not?
3. Design an open-ended, but well-planned activity in which students in pairs or groups carry out a specific task to their liking or interest (e.g., summarizing a book chapter, evaluating a course textbook, convincing a partner to invest in a certain company, making a hotel reservation on the telephone, etc.). On completion of the task, ask students to report to the class on their findings individually or collectively, orally, and / or in writing. After that, ask them to self-evaluate and / or other-evaluate their performance using specific rubrics or criteria. Survey your students on how much they feel they have benefited from the task itself as well as from self-evaluation and /or other-evaluation.
4. Compare the performance and achievements of students in a traditional language-centered, teacher-centered class with those of students in a learner-centered TBLT and TBLA class. Do you see any differences? Specify the main difference you have observed between the two contexts. Discuss the major differences you have observed with your colleagues or students.

References

- Andrade, H., & Cizek, G. (Eds.) (2010). *Handbook of formative assessment*. New York: Routledge.
- Black, P., & Wiliam, D. (2003). In praise of educational research: Formative assessment. *British Educational Research Journal*, 29(5), pp. 623–637.
- Brindley, G. (1994). Task-centred assessment in language learning programs: The promise and the challenge. In N. Bird, P. Falvey, A. Tsui, D. Allison, A. & McNeill (Eds.), *Language and learning* (pp. 73–94). Hong Kong: Institute of Language in Education.
- Brown, H. D. (2004). *Language assessment: Principles and classroom practice*. White Plains, NY: Pearson ESL.
- Carless, D. (forthcoming). TBLT in EFL settings: Looking back and moving forward. In A. Shehadeh & C. Coombe (Eds.), *Researching and implementing task-based language learning and teaching in EFL contexts*. Amsterdam: John Benjamin's.
- Cheng, L. (2008). The key to success: English language testing in China. *Language Testing* 25(1), pp. 15–37.
- Ellis, R. (2000). Task-based research and language pedagogy. *Language Teaching Research*, 4, pp. 193–220.
- Ellis, R. (2003). *Task-based language learning and teaching*. Oxford: Oxford University Press.
- Gan, Z., Davison, C., & Hamp-Lyons, L. (2009). Topic negotiation in peer group oral assessment situations: A conversation analytic approach. *Applied Linguistics*, 30(3), pp. 315–334.

- Long, M. H., & Norris, J. M. (2000). Task-based language teaching and assessment. In M. Byram (Ed.), *Encyclopedia of language teaching* (pp. 597–603). London: Routledge.
- Luk, J. (2010). Talking to score: Impression management in L2 oral assessment and the co-construction of a test discourse genre. *Language Assessment Quarterly*, 7(1), pp. 25–53.
- Ng, C., & Tang, E. (1997). Teachers' needs in the process of EFL reform in China: A report from Shanghai. *Perspective* 9, pp. 63–85.
- Poehner, M. (2008). *Dynamic assessment: A Vygotskian approach to understanding and promoting L2 development*. Berlin: Springer.
- Richards, J., Schmidt, R., Platt, H., & Schmidt, M. (2003). *Longman dictionary of applied linguistics*. London: Longman.
- Robinson, P., & Ross, S. (1996). The development of task-based assessment in English for academic purpose programs. *Applied Linguistics*, 17, pp. 455–476.
- Shehadeh, A. (2005). Task-based language learning and teaching: Theories and applications. In C. Edwards & J. Willis (Eds.), *Teachers exploring tasks in English language teaching* (pp. 13–30). London: Palgrave Macmillan.
- Shehadeh, A., & Coombe, C. (Eds.) (2010). *Applications of task-based learning in TESOL*. Alexandria, VA: TESOL.
- Shehadeh, A., & Coombe, C. (Eds.) (forthcoming, 2012). *Researching and implementing task-based language learning and teaching in EFL contexts*. Amsterdam: John Benjamin's.
- Van den Branden, K., Bygate, M., & Norris, J. (Eds.) (2009). *Task-based language teaching: A reader*. Amsterdam: John Benjamin's.
- Weaver, C. (forthcoming). Incorporating a formative assessment cycle into task-based language teaching. In A. Shehadeh & C. Coombe (Eds.), *Researching and implementing task-based language learning and teaching in EFL contexts*. Amsterdam: John Benjamin's.
- Wigglesworth, G. (2001). Influences on performance in task-based oral assessments. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching, and testing* (pp. 186–209). Harlow, England: Longman.
- Winke, P. (2010). Using online tasks for formative language assessment. In A. Shehadeh & C. Coombe (Eds.), *Applications of task-based learning in TESOL* (pp. 173–185). Alexandria, VA: TESOL.
- Yeh, A. (2010). Assessing task-based activity in a speech training class. In A. Shehadeh & C. Coombe (Eds.), *Applications of task-based learning in TESOL* (pp. 161–171). Alexandria, VA: TESOL.