Test Modification Project

Komila Sayfiyeva Kamolidinovna

English teacher at Tashkent State Law University, Tashkent, Uzbekistan

Aygul Aghabalayeva

English teacher at Azerbaijan University of Languages, Baku, Azerbaijan

Annotation: Test has significant influence on learning as well as teaching process. Assessing students' performance is one of the most critical aspects of a teacher (Mertler, 2009). For teachers it is crucial to use a test so as to know and measure whether the objectives of language learning and teaching have been achieved or not. The current Test Modification Project is aimed to investigate one of the existing tests and find out the flaws of the test as well as modify it in terms of five key principles. The importance of the project for English language teachers is that it gives a chance to analyze assessment tools by putting their theoretical knowledge into practice and to adapt them according to the needs of their learners.

Keywords: The test is summative assessment tool, achievement test, formative assessment, modified version, validity, practicality of the test, authentic test, wash back, original version, content-related validity issues.

INTRODUCTION

Language assessment and testing is considered challenging part of teaching, as it requires careful analysis before administrating any kind of test or assessment. However, this process is positive in terms of progress that learners are making during their studies. Additionally, language testing and evaluation have long been seen as the most important and appealing aspects of language instruction and acquisition. The reason for this is that, these methods can not only be used to determine a student's development in their knowledge by administering suitable examinations, assessing the findings, and providing feedback, but they can also be used to improve the learners' achievement. A good teacher should find appropriate way of assessing their students based on their teaching methods, approaches or techniques, also the needs of the learners. Brown defines assessment as "an ongoing activity encompassing a wide range of methodological tools." So, rather than simply grading students based on their information, the major objective of an assessment should be to provide them the opportunity to present the skills they developed over a period of time. A good teacher's job is to select the appropriate form of assessment, recognize its significance, figure out how to demonstrate it, and analyze the assessment system's results. This project analyzes one assessment tool used in the class as a summative assessment, and it explores some modifications namely, some changes, recommendations and outcomes, as well as future implications of the very test.

METHODS

A description of the learner

Although there were many possible candidates who are suitable for conducting the project, we had difficulties in choosing one of them. After considering many aspects, we have decided to fulfill the project with one of the intermediate learners who have recently finished pre-intermediate course.

The initial reason is that her strong desire and determination of learning English. She wants to enter World Languages University no matter what although she has failed the entrance exam twice. Another factor which made us chose her is despite the fact that she is knowledgeable and always prepares home tasks without any flaws, when it comes to exams she tends to lose herself which in turn affects her results negatively. Last but not least, there is usually inconsistency between her language ability and test results which makes her dissatisfied about her test results. When we told her about the project she was eager to attend and agreed willingly to collaborate with us for the accomplishment of the project (Appendix 1). Before analyzing her results, we would like to give information about general background of the test taker since it is crucial to know.

Face to face interview with the target student was essential to collect the necessary data about her. Mehriniso is a 20 year old girl who is currently preparing only for entrance exams of the university. She is bilingual learner whose native language is Uzbek and second language is English. Although her parents are monolingual, they supported and constantly gave her motivation to learn languages. She started to learn English at the age of 17 and she told that she learnt only grammar till she reached 18. After that although her level was elementary she also started to prepare CEFR exams with the teacher. Studying the course of CEFR was a real challenge for her and for this reason she decided to leave that center and opted for my classes. When she first came in our language center, we found out that she was at elementary level through placement test. As Brown (2010) emphasized one of the benefits of placement test is to get diagnostic information on a student's performance, which helps teacher to assess her students' abilities from the very start. The placement test showed the following results:

Table 1. Results of first placement test

	Overall Questions	Results
Grammar	20	10
Vocabulary	20	8
Reading	20	7
Listening	20	14
Speaking	20	10
Overall	100	49

It is also noteworthy to say that, according to the student, when she sits an exam all the knowledge suddenly disappears out of her mind because of the fear of failure. We thought that this was the main reason why she has inconsistent results. To specify, she got 93 out of 100 from the final exam of elementary level whereas in the final pre-intermediate exam she scored 71 out of 100.

A description of the most recent Placement Test

Mehriniso started to study in pre-IELTS group after taking final pre-intermediate examination. Initially, We wanted her to finish intermediate level first and then to start preparing for IELTS exam. However, she was eager to get the certificate of IELTS as soon as possible as she had few months left for the entrance exams. That is why, we decided to conduct placement test whether she is ready for studying IELTS lessons or not and to place her into the right group. The placement test was totally different from the one which we usually conduct for the new comers. The test includes four skills along with grammar and vocabulary. We decided not to include IELTS materials as they are very difficult for students' level. 20 students from my pre- intermediate and intermediate courses were involved too. Overall, there were 50 questions, 10 questions for each section. Since there were many items in the test, it took 3 hours to conduct the test. As a result, the students got exhausted at the end. This time Mehriniso surprised me with her good results:

Table 2. The results of the recent placement test

Student	Listening (100)	Reading (100)	Writing (100)	Speaking (100)	Grammar &Vocabulary (100)	Overall
Raimova Mehriniso	80	72	78	85	91	81

Overall, 11 students could pass the exam and placed in one group. The target student also passed the exam with 81 score which shows that she is ready to start IELTS lessons. After the test, each student got appropriate constructive feedback to work on their weak side.

CRITIQUE OF AN EXISTING LANGUAGE TEST

Description of the listening test

The listening section of the final pre-intermediate test was chosen to analyze for this project because the target student had inconsistent results which in turn made her dissatisfied with her listening score. The test is summative assessment tool, which students take at the end of the pre-intermediate course. As for the purpose of the test, it is an achievement test since its main objective was to check whether the students gained the knowledge which was provided during the course or not. Hughes (2003) points out that achievement tests are directly related to language courses, their aims being to identify how successful was individual students, groups of students or courses themselves have been in order to achieve objectives.

ORIGINAL VERSION OF THE LISTENING TEST:

As for General English lessons, typically four integrated skills are taught together in our learning center, the final test also aims to assess reading, listening, writing and speaking. The final test was taken from the course book which is called "Headway". Overall, there are 10 items in listening section which consists of 5 multiple choice and 5 matching questions. For each correct response in multiple choice questions is given 1 point and 2 points for matching questions; 25 points in total. The audio lasts about 4 minutes and it consist of monologues of 5 speakers about their work. The given features of the test depict that it is traditional assessment. Considering the fact that the listening test is neither too long nor too short, all the items will be adapted and some necessary items will be modified taking into account of the needs of the learner as well as the strength and weaknesses of the test.

According to Brown (2004), tests should require some specific criteria for testing which include practicality, reliability, authenticity, validity and wash back. The analysis of the listening section based on the very principles of language assessment gave me opportunity to know the reasons why the results did not live up my and the target learner's expectations.

RESULTS AND ANALYSIS

Analysis of the listening test according to 5 principles

A Practical test is easy to administer, has a scoring procedure which is specific and at the same time, considers the time and effort involved for both design and scoring (Brown, 2010). The test is considered practical for the teacher as it is easy to administer because the test consists of matching and multiple choice questions. The test also consists of clear scoring procedure which is time efficient. The only thing teachers need to do is to count correct answers.

The test under the study has some problems related to the principles of **reliability.** Brown (2010) suggested that a reliable test should have clear instructions and uniform rubrics for assessing students' performance and if we conduct the test many times, the results should be the same. The instructions are not clear and at the beginning it was not even written or said what students are going to listen about. Furthermore, the test does not have any examples which make the task unambiguous. The chosen question types are not good enough to determine the real knowledge of the test-takers but encourages them guessing. For example, matching and multiple choice questions which caused the target test-taker to guess and exhausted by the time she reached the next test item. So, there is a student-related reliability issue. However, the listening test is reliable in terms of rater reliability factors since there is given clear scoring criteria and does not cause any bias while scoring.

While **validity** is the most essential principle of any assessment, the main problems of the majority of the tests are observed in the validity (Popham, 2008). It is important for the test to be valid so that the results of the test will be accurately applied and interpreted. The listening test is found to be invalid for many reasons. The purpose of the test is to assess students' comprehension and the items which were chosen for this are inappropriate, hence it lacks construct validity since students try mostly guess and find correct answers rather than to understand the meaning of the monologues. The test is also lacked in content-validity since the vocabulary and topic were not previously covered in the lessons. The learner mentioned that some items were inappropriate, too difficult or too easy. The issues with content and construct related validity of the test exerted a direct influence on the face validity of the test. To be more specific, the target learner was discouraged with her results.

With respect to **authenticity,** Bachman and Palmer (1996), "Authenticity is the degree of correspondence of the characteristics of a given language test task to the feature of the target

language task". In other words, authentic test should contain real-world tasks, natural use of the language along with meaningful, appropriate topics. Listening task includes real-world monologues of five native speakers about their job. Nevertheless, since the students were not familiar with the topic and the vocabulary in previous lessons it shows lack of authenticity. Moreover, although it manifests natural language but the tasks cannot stimulate a student's performance in reality by providing only multiple-choice or matching testing items.

According to Brown (2010), **washback** improves several basic principles of language acquisition such as self-confidence, language ego, intrinsic motivation as well as autonomy of learning. Since the feedback for students was not designed for the language development of the test taker, but only gives numerical scores, it does not provide any positive or negative washback for the learner.

A critique of the listening test

Strengths	Weaknesses
➢ The organization of the test is accurate and relevant because test items are logically sequenced from easy to difficult questions ➢ The test is very practical as it is easy to conduct, score and does not take much time for students to solve and for teachers to check ➢ Each listening task shows how many points will be given for correct responses which makes the learner aware about how the answers will be scored	➢ Clear instructions and examples for each task are not given ➢ The test lacks to measure the students' true comprehension and ability of listening as it contains only matching and multiple choice questions with a lot of options which means, there is high possibility of guessing. ➢ The topic and vocabulary of the listening text were not covered during the lessons.

Taking into consideration the strengths and weaknesses of the listening tests, I am going to change and adapt it according to my learner's need and preferences. Firstly, pre, while and post listening activities will be created to make the learner aware of the topic and vocabulary as well as to enhance the acquisition of the language learning process. Secondly, test items will be modified into other questions types with the reduction of some options. Lastly, the clear instructions with appropriate examples will be provided.

MODIFIED VERSION OF THE LISTENING TEST

Suggestions for improvement

The main aim of listening skill is to enable the learner to perceive L2 like native speakers (Robinett, 1978). This is the ultimate goal which teachers want to achieve with their students and it depends on mostly the tests which they take during the lesson. While creating the test for listening, it is crucial for the teacher to take into account of the level of the students along with their needs, the speed of the listening, authenticity, instructions and chosen test items. It is also vital to pre-teach before implementing one particular task in the classroom. As Morley (2001) suggested that students must always know exactly what the task is, such as what to listen for, where to listen, when and how to listen.

The following suggestions would be necessary for the development of the listening test:

1. As the content and vocabulary were not familiar to my learner, the list of complicated words which were used in listening activity are going to be provided and Vocabulary quiz will be created so as to foster the comprehension of the test. The quiz is going to be conducted online with the help of www.quizizz.com website in order to make the process of doing and checking easily and quickly.

2.   The listening test should also carry itself positive washback and teachers are responsible for creating friendly atmosphere by including pre-listening and post listening activities. According to my learner, the listening test was out of nowhere and initially she had no idea what she was listening.

3.   The number of test items is going to be increased and some test items will be changed to challenge the strengths and weaknesses of my learner. For example, there were many options in matching questions which caused the learner to confuse. Some questions should be modified into short-answer questions or sentence completion tasks.

4.   The instructions with clear examples need to be also added to create comfortable testing environment. To clarify, in the first task of the listening, it only says, *listen to five people and circle the correct letter.* In addition, there is not given any example in the second task of the test.

Taking everything into consideration, I tried to modify the listening section of the existing summative test by paying attention to the flaws of the test based on the five principles and administer the new version of the assessment tool to the test-taker to identify the differences between original and modified listening test.

MODIFIED VERSION OF THE TEST

Instructions of the listening test of final pre-intermediate exam for students

You are going to listen to five people talking about their work. Since the related vocabulary was not covered during the lessons, you will be provided with the list of words with their definitions (Appendix 2). The test consists of 3 parts and lasts 20 minutes:

1. PRE-LISTENING STAGE.

You will look through the words with definitions in 4 minutes and after that do online vocabulary quiz in the website www.quizziz.com in 3 minutes (Appendix 2)

2. WHILE LISTENING STAGE.

I.   Listen to 5 people talking about their work and answer the questions *no more than three words or a number.*

1.   When did Fiona buy baker's shop? *5 years ago*

2.   Where does Fiona work?

3.   Which work was stressful for Jason?

4.   Which language Peter could not speak?

5.   Why did Donna give up cycling?

6.   Who is Chris jealous of? *2 points for each correct answer* | 10 |

II. Listen again and choose your answers from the box and write the correct letters next to questions 1-6. *Who matches each description?*

F   Fiona
J   Jason
P   Peter
D   Donna
C   Chris

1. does not earn much money but doesn't mind.   *J*
2.   sometimes feels envious about other people's work.   …………….
3.   helps other people get over their difficulties   ……………
4.   teaches English at his/her own school.   ……………
5.   enjoys working on his/her own.   …………..
6.   hired two people to work for him/her.   ………….

1 point for each correct answer | 5 |

III. Post listening activity.

You are required to write a short essay about **"My parents' work"** by using the words below. You should use at least 3 of them. You will have only 5 minutes to finish the task.

Earn a living, day off, eat out, proper, hire, satisfied with

_____

_____

_____

_____

_____

_____

5 points for thefull answer | 5 |

Alignment of modified version with ESL standards at the proficiency level of the student

It is clear that the modified version of the listening test is more effective as the format is very accurate which in turn leads the test-taker to follow easily and teacher to administer and assess efficiently:

➢ Instructions and examples for each task were specified and given clearly to the teacher and the students which can cause the efficiency of the test. Additionally, the time limit for each task was clearly given so as to manage their time effectively.

➢ There is a clear direction for scoring which can be calculated easily and does not cause the fatigue or bias for the teacher to evaluate since there is high reliability in terms of chosen test items. Students also do not get exhausted or confused by answering the questions as there are not many options. Furthermore, the pre, while and post listening tasks help the student to reduce anxiety during the test. Student-related reliability as well as rater reliability was addressed in this test.

➢ Content-related validity issues were also solved with the help of the pre-listening task which involves giving definitions for each unfamiliar words and online quiz helps to memorize the words. The test intends to measure the listening comprehension of the learner and the questions are designed only for this purpose.

➢ The test can be considered as an authentic test since in the last part of the test, students are required to write an essay by using the words which they may encounter in real-life situations and the topic of the short essay matches with the real-world tasks.

➢ The aim of the last part of the test is to provide a positive washback for the students and students can have a chance of getting constructive feedback based on their written performance.

➢ The questions are appropriately designed for the level of the students as the curriculum of the course includes B1+ level topics and vocabulary.

➢ By the help of Online Vocabulary Quiz and the list of words provided before the listening students have a chance to revise and memorize the new words which were used in the listening.

DISCUSSION

Administration of the modified test

After introducing some modifications to the listening test, I administrated new version of the test to my student. The administration of the test consisted of two parts which are internet-based where the participant needed to do online quiz and paper-based. For both, clear guidelines for the test-taker were given. Before conducting the listening test, the voice of the audio player were checked thoroughly and temperature, light and other conditions of the room were also taken into account in order to prevent test administration unreliability and provide the best testing atmosphere for the learner.

Comparison of results of original and modified version of the test

Regarding original version of the test, Mehriniso could find only 9 correct answers out of 25 questions. On the other hand, she scored 18 out of 20 in the modified version (Appendix 3) which could live up my expectations. From my observations, she did not encounter any difficulties while taking the test since the instructions along with examples were provided clearly and appropriate time for each task was given as well. Furthermore, pre-listening task prepared her for the while listening task, post listening task facilitated her comprehension more deeply. When the participant was asked her opinions about two tests, she responded that the first test made her worry a lot and she was not prepared well for the listening since it took time for her to grasp what are the speaker are talking about and as a result, she ended up guessing the answers. She also mentioned that as the vocabulary was unfamiliar for her, she had difficulties to catch the meaning of the speeches. As for second test, she said the stages of listening tasks were really helpful for her to comprehend the listening. She admitted that the instructions were clear enough and she enjoyed doing the vocabulary quiz online since it exerted her further understanding. Moreover, the type of questions which were selected was very useful and appropriate for her level.

Overall, the principles of testing have met in the modified version in many ways. For example, validity and reliability issues were tackled by the means of pre-listening and while- listening activities. Authenticity can be seen in the test items which are correlated with world- tasks whereas washback mostly have been successful in post-listening task. However, in terms of practicality, although to check and administer the test is easy, only designing of the test can be a hindrance to the teacher as it takes much time.

Conclusion

Rationale and Justification

The listening section of the final pre-intermediate exam was chosen and modified for this project for many reasons. Firstly, not only the test-taker, but many students had inconsistent and unsatisfied results from the listening. Secondly, there were many flaws in the listening section and

inappropriate question items were chosen. While conducting this project, I realized that it is vital for teachers to understand and implement the principles of language assessment in evaluating the existing assessment procedure, or designing a new one by themselves (Yoneda, 2012). This project also helped me to do further research about the five main principles and how to apply them in modifying the test items. Although I have been teaching for 3 years, I have never paid attention to the importance of assessment tools in teaching and learning process. Admittedly, I have administrated the given test provided by the authorities of our educational center with my students without considering validity, reliability, authenticity and washback since as Stiggins (2002) pointed out that teachers rarely have the opportunity to learn how to use assessment as a teaching and learning tool. That is why, I am very grateful that I had a chance to put into practice what I have learned theoretically during the sessions.

Recommendations

Regarding recommendations, it would be very efficient if the test modified fully, but not only one section. During the project and studying the test carefully, I realized that there were also some deficiencies in terms of other skills as well. For example, not all skills were integrated and there were lack of content validity, reliability and authenticity problems too. Moreover, I had an opportunity for studying and testing only one student thoroughly that is why, the test also needs to be piloted with whole group of the learners in order to be sure the effectiveness of the new version of the test. Additionally, while conducting the very test every teacher needs to be careful and study the test in detail as the modified version of the test was adapted according to my students' needs and preferences.

Future implications

As this new version of the test was successful with my learner, in the near future, I am going to propose the very project to the authorities of "Talent Development Center". Since there were much dissatisfaction among students and the director of the center, I hope this new version of the listening test would promise good outcomes. Besides that, I am going to pilot modified version with my students who are studying online these days. Furthermore, from now on, when conducting the test or examination I try to be more considerate in assessing students as Cameron (2001) stated that assessments can inspire students, help teachers to conduct effective lessons, support further learning with feedback.

References:

1. Bachman, L., & Palmer, A. S. (1996). Language testing in practice. New York: Oxford University Press.

2. Brown, H. D. (2004). Language Assessment: Principles and Classroom Practices. New York: Pearson Education.

3. Brown, H. D. & Abeywickrama, P. (2010). *Language assessment, principles and classroom practices* (2nd ed.). White Plains, NY: Pearson Education.

4. Cameron, L. (2001). Teaching languages to young learners. Cambridge, UK: Cambridge University Press.

5. Hughes, A. (2003). *Testing for Language Teachers* (2nd ed.). Cambridge England: Cambridge University Press.

6. Kamolidinovna, S. K. (2021). MODERN AND INTERACTIVE METHODS OF TEACHING LEGAL ENGLISH. *CURRENT RESEARCH JOURNAL OF PHILOLOGICAL SCIENCES (2767-3758)*, *2*(12), 190-194.

7. Mertler, C. A. (2009). Teachers" Assessment Knowledge and Their Perception of the Impact of Classroom Assessment Professional Development. Improving Schools, 12 (2), 101-113.

8. Morley, J. (2001). Aural comprehension instruction: Principles and practices. In M. Celce-Murcia (Ed.), Teaching English as a second or foreign language (3rd ed.). Boston: Heinle & Heinle

9. Popham, W. J. (1997). What's wrong—and what's right—with rubrics. Educational Leadership, 55, 7 2 -75.

10. Robinett, B.W. (1978). Teaching English to speakers of other languages: Substance and technique. New York: McGraw-Hill Book Company.

11. Sayfiyeva, K. K. (2021). CHALLENGES AND METHODS OF TEACHING ENGLISH TO LAW STUDENTS. *Eurasian Journal of Academic Research*, *1*(9), 610–614. извлечено от https://www.in-academy.uz/index.php/ejar/article/view/236

12. Sarvara T. Akobirova, Komila K. Sayfiyeva. (2022). Importance of Learning Styles: How does Students' Awareness of Learning Styles affect their Proficiency in Second Language Acquisition (SLA). *International Journal of Early Childhood Special Education (INT-JECSE),* 14(1): 802-807

13. S. Seytniyazova, 2021. Teaching English through interactive games in a higher educational establishment. Eurasian Journal of Academic Research 1 (9), pp 584-588.

14. Stiggins, R. J. (2002). Assessment Crisis: The Absence of Assessment for Learning. Phi Delta Kappan, 83(10), 758-765.

15. Yoneda, M. (2012). Designing Assessment Tools: The Principles of Language Assessment.

16. Bull. Mukogawa Women's Univ. Humanities and Social Sci., 60, pp. 41-49