



MODIFYING READY-MADE AND TIMID TESTS INTO RELIABLE , AUTHENTIC AND PRACTICAL TEST IN EVALUATING STUDENTS' KNOWLEDGE AT SCHOOL

Journal Website:
<https://masterjournals.com/index.php/crjps>

Copyright: Original content from this work may be used under the terms of the creative commons attributes 4.0 licence.

Submission Date: February 18, 2023, Accepted Date: February 23, 2023,

Published Date: February 28, 2023

Crossref doi: <https://doi.org/10.37547/philological-crjps-04-02-05>

Nozima Siddikova

EFL/ESL Instructor Uzbek State World Languages University Tashkent, Uzbekistan

ABSTRACT

In the following article, I try to show some drawbacks and lacks of ready-made test which has been taken at school to identify students' level. Also, I put modified version of test instead of the timid test at school taking into consideration five assessment principles. Moreover, I present student's before and after result depending on the administered tests.

KEYWORDS

Knowledge at school, present student's, result depending.

INTRODUCTION

It is true that these days majority of practicing teachers are likely to use ready-made and simple tests to measure their students' language knowledge. We have to admit as a teacher that it is a kind of teachers' habit, which would be the sign of unwillingness to work on assessment. As a

result, students may scare of tests or feel tense up when their teacher mentions about a test. The reason for that is ready-made and timid tests cannot evaluate test-takers' performance objectively because of invalidity, unpracticality or negative wash back. In the following project, I will inform about midterm test evaluating the



grammar and vocabulary skills, which has been taken recently and discuss its strengths and weaknesses along with modification version.

Learner assessment profile

As I had worked as an EFL teacher at school №26 located in Zangiota district, I have chosen one of my student's assignment to analyze and modify because of her consent. When I informed pupils about the given test modification project, only Mahfuza Mirqosimova agreed her test result to be analyzed in order to identify the strengths and weaknesses of the multiple choice question test that focus on grammatical and vocabulary skills of students. Mahfuza Mirqosimova, who is 15 years old, studies at 8th grade. Mahfuza Mirqosimova has been studied English since 5th grade. The taken test result shows that her level was A2+ because she found 20 correct answers out of 25 questions. While I am conducting lesson, I seek to observe all my pupils performance during process. According to my observation, Mahfuza Mirqosimova looks active and interested in language learning, specifically, she is able to speak in English and Russian languages. However, her result made me a bit hesitated in measuring her real knowledge because mainly she made mistakes on grammar questions. Consequently, it cause to analyze the test in depth, and transform into integrative and extended type of assessments.

The educational setting

Although the school №26 is situated in Zangiota district, Tashkent, it is one of the well-developed public schools. 45 teachers work as practicing teachers that most of them have been awarded with many certificates for their dedication to their job and specialties. According to curriculum, English lessons are conducted three times a week and each lesson lasts 45 minutes as all public school systems. Additionally, the school has been equipped with the all facilities and technological devices for teaching and learning process based on the use of advanced information and media technologies in the Presidential Decree №1875.

The most recent placement test

During the teaching and learning process, I have to evaluate pupils' knowledge what they learned in each lesson based on formative type of assessment such as giving feedback or peer evaluation, as Tomlinson (2001) states that formative assessment is more important rather than summative assessment. Therefore, teachers should approach the individually, despite that they assess their students comparing to each other. However, the following schools placement test has been taken as a summative assessment as it is a midterm test. Brown (2010) states that multiple choice tasks are easy to administer and score but they are difficult to create. The taken multiple-choice task, which is ready-made tests, consists of 25 selected response questions. Furthermore, it focuses on testing knowledge rather than language knowledge as more questions are based



on 8th grade textbook. Mahfuza got 5 mark as she found 20 true answers out of 25.

Critique of an assessment

The following test is a ready-made test taken from a test bank for the midterm according to the curriculum of the school. It was aimed to check pupils' grammar and vocabulary knowledge but mostly it concerns on the informative knowledge learned contexts in their textbook material during the lesson. A good test should be based on five principles of assessment. However, in my case, there are some issues on those principles such as validity, reliability and wash back.

Strengths of test

Practicality

According to Brown (2010) practical test stays within budgetary limits and be finished by the test taker within proper time constraints. Besides, it has clear directions for administration considering the time and effort involved for both design and scoring. Therefore, one of the strengths of the test is its practicality.

Weaknesses of the original test

Validity

This multiple-choice test is not face valid because some of the tests focus on knowledge, rather than language skills. A test is said to have high validity when the test accurately measures what is intended to measure (Hughes, 1993).

Here, you can see only two questions on grammar skill out of ten questions. It should have focused on more grammar and vocabulary. However, it will not be mistake if I say that this test is culturally and linguistically appropriate for the learner because all the things in the test were taught to Mahfuza Mirqosimova during the lesson. As administering the test and scoring are uncomplicated, I preferred to administer this test.

Reliability

A reliable test, according to Bynom (2001), will contain well-formulated tasks and not indefinite questions; the student will know what exactly should be done. The test will always present ready examples at the beginning of each task to clarify what should be done. However, in the original test was not given any example. Moreover, Heaton (1990), states that the test could be unreliable if the two different markers mark it. For example, one representative of marking team can be more lenient and have different demands and requirements, but the other one could appear to be too strict and would pay attention to any detail. In terms of reliability, multiple-choice test is not reliable considering no exact rubric that may cause fluctuation in scoring. Therefore, I witnessed some differences between marks put by two teachers, as they approached differently due to no rubric.

Authenticity



According to Bachman (1991) authenticity is considered as the appropriateness of a language user's response to language as communication. The test items should be related to the target language's use. Multiple-choice tasks are sometimes viewed as not being authentic language use (H. Brown, 2010). The original test is not authentic, it does not ask to show the any skill because Wiggins (1993) believes that authentic test includes a task for pupils to perform specific skill and a rubric by which their on the task will be assessed.

Wash back

Brown (2010) mentions formal tests provide no beneficial wash back if the students receive a simple grade or a single overall numerical score. An original test is used as a mid-term test. Mid-term tests are a type of formal formative tests. To score, scale testing is used: overall numerical score at the end. If pupil has a chance to get feedback from teachers, they will have positive wash back.

Modified version of the chosen test

Geoff (2015) good teaching involves assessment that identifies the learner needs and design of tests and learning activities that challenge and extend them. Before administering a test, teacher have to be aware of needs of learner. Based on my experience, I felt that Mahfuza Mirqosimova needed to develop her communicative skills along with grammar. Having listened her comments on

the original test, I had to modify the current test into the task that focus on checking grammar and vocabulary skills through speaking.

More precisely, I changed this test according to five principles so that it could be good featured test.

Instruction: You have five minutes to think about the topic and 10 minutes to make up questionnaire about the films that you watched. After writing questionnaire. Talk with your partner to survey each other for 2 minutes taking turn.

Task: Talk about the films and actors that you watched by answering following questions.

In your questionnaire, you should include:

- What was it?
- When was it?
- Where was it?
- Who played roles?
- What happened?
- Why it is the most favorite?

You may use the following words in your speech: Gulchekhira Jamilova, "Past Times", actor, goodies, baddies, characters, funny, comedy, detective.

Authenticity

Firstly, in terms of authenticity, I modified selected response type of test into extended response task. Regarding Bachman and Palmer (1996), the purpose of extended production is to obtain larger amounts of language from test-taker and to



allow for more construction that is creative; therefore, these tasks are likely to elicit instances of authentic language use. Learner can speak about the film that she watched in the real life.

Reliability

Secondly, I used speaking rubric to evaluate the learner by several criteria. A rubric is a coherent

set of criteria for students' work that involves descriptions of levels of performance quality on the criteria and the main purpose of rubrics is to assess performances. A reliable test has uniform rubrics for scoring/evaluation.

	Fluency	Pronunciation and accent	Vocabulary	Grammar	Details
5	Smooth and fluid speech; few to no hesitations; no attempts to search for words; volume is excellent.	Pronunciation is excellent; good effort at accent	Excellent control of language features; a wide range of well-chosen vocabulary	Accuracy & variety of grammatical structures	Excellent level of description; additional details beyond the required
4	Smooth and fluid speech; few hesitations; a slight search for words; inaudible word or two.	Pronunciation is good; good effort at accent	Good language control; good range of relatively well-chosen vocabulary	Some errors in grammatical structures possibly caused by attempt to include a variety.	Good level of description; all required information included
3	Speech is relatively smooth; some hesitation and unevenness caused by rephrasing and searching for words; volume wavers.	Pronunciation is good; Some effort at accent, but is definitely non-native	Adequate language control; vocabulary range is lacking	Frequent grammatical errors that do not obscure meaning; little variety in structures	Adequate description; some additional details should be provided
2	Speech is frequently hesitant with some sentences left uncompleted; volume very soft.	Pronunciation is okay; No effort towards a native accent	Weak language control; basic vocabulary choice with some words clearly lacking	Frequent grammatical errors even in simple structures that at times obscure meaning.	Description lacks some critical details that make it difficult for the listener to understand
1	Speech is slow, hesitant & strained except for short memorized phrases; difficult to perceive continuity in speech; inaudible.	Pronunciation is lacking and hard to understand; No effort towards a native accent	Weak language control; vocabulary that is used does not match the task	Frequent grammatical errors even in simple structures; meaning is obscured.	Description is so lacking that the listener cannot understand

Thirdly, discrete-point test was changed into integrate-point test. Integrating the language skills also promotes the learning of real content, not just the dissection of language forms. Finally, the integrated-skill approach, whether found in content-based or task-based language instruction or some hybrid form, can be highly motivating to students of all ages and backgrounds Validity.

The next change is that I set time limit (17 (5+10+2) minutes), instruction, and key vocabulary for the

task. According to five principles, teachers can enhance student`s perception of fair tests by using

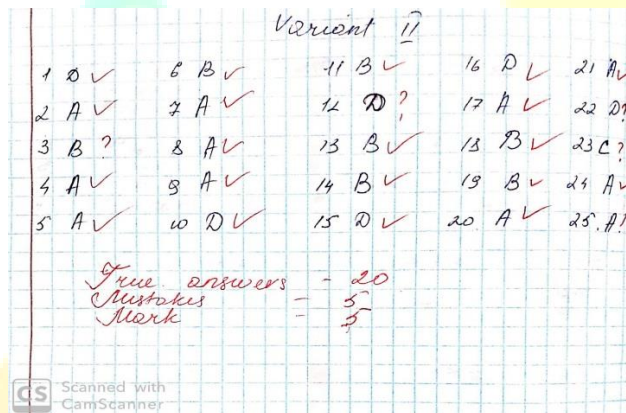
- Well-constructed, expected format with familiar tasks
- Tasks that can be accomplished within an allotted time limit
- Items that are clear and uncomplicated
- Directions that are clear
- A difficulty level that presents a reasonable challenge

Wash back



As I inserted rubric form to measure learner's knowledge, Mahfuza had positive wash back because of the opportunity to be aware of exact mistakes and working on those mistakes through getting feedback from teacher. What is more, I could approach the learner individually. It is clear that it helps to assess learners' real ability in EFL. Piloting the modified version of the test

When I administered the modified version of test to the Mahfuza Mirqosimova, she did not have much difficulty in the testing process. The reason for that was I did not change the context of the original test. As a result, learner could easily speak about the topic using the given key words after thinking for 2 minutes. Furthermore, another surprising point was the difference between the results in the original and modified version of tests. In the original test, Mahfuza got 5 while she got 4 in the modified test.





Muslimo

	Fluency	Pronunciation and accent	Vocabulary	Grammar	Details
5	Smooth and fluid speech; low to no hesitations; no attempts to search for words; volume is excellent.	Pronunciation is excellent; good effort at accent.	Excellent control of language features; a wide range of well-chosen vocabulary	Accuracy & variety of grammatical structures ✓	Excellent level of description; additional details beyond the required
4	Smooth and fluid speech; few hesitations; a slight search for words; inaudible word or two ✓	Pronunciation is good; good effort at accent. ✓	Good language control; good range of relatively well-chosen vocabulary	Some errors in grammatical structures possibly caused by attempt to include a variety.	Good level of description; all required information included
3	Speech is relatively smooth; some hesitation and unevenness caused by rephrasing and searching for words; volume wavers.	Pronunciation is good; Some effort at accent, but is definitely non-native	Adequate language control; vocabulary range is lacking ✓	Frequent grammatical errors that do not obscure meaning; little variety in structures	Adequate description; some additional details should be provided ✓
2	Speech is frequently hesitant with some sentences left uncompleted; volume very soft.	Pronunciation is okay; No effort towards a native accent	Weak language control; basic vocabulary choice with some words clearly lacking	Frequent grammatical errors even in simple structures that at times obscure meaning.	Description lacks some critical details that make it difficult for the listener to understand
1	Speech is slow, hesitant & strained except for short memorized phrases; difficult to perceive continuity in speech; inaudible.	Pronunciation is lacking and hard to understand; No effort towards a native accent	Weak language control; vocabulary that is used does not match the task	Frequent grammatical errors even in simple structures; meaning is obscured.	Description is so lacking that the listener cannot understand

CONCLUSION

To conclude, I think that the all factors that I learned by changing the existing test were immensely important in my future teaching process. Here, I acquired the huge knowledge about testing tools and put them into practice to check students' grammar through integrated skills. In addition, I used five principles of testing and assessment in designing a test. I think that this modified integrated-point and extended response test will be more reliable and valid attaching rubric in teaching process. It gives a chance for me to provide my learners with appropriate feedback individually. As a result, they can be aware of their mistakes and work on correcting those mistakes. Moreover, testing the pupils the teachers should be aware of the fact that together with checking

their knowledge the test can adversely affect pupils. Therefore, the teachers have to design such a test that it could encourage the students, but not to make them reassure in their own abilities. The test should be a friend, not an enemy. Thus, the issue of validity and reliability is very vital in creating a good test. The test should measure what it is supposed to measure, but not the knowledge beyond the students' abilities. Moreover, the test will be a true indicator whether the learning process and the teacher's work is effective.

REFERENCE

1. Brown.H.D. (2010). Language assessment principles and classroom practices. Pearson



Education, 10 Bank Street, White Plains, NY 10606

2. Geoff T (2015). Putting students to the test . Issue Twenty. Forum. July
3. Heaton, J. 1990. Classroom Testing. Longman
4. Hughes, A. (1993). Testing for language teachers. Cambridge, MA: Cambridge University Press
5. Mousavi, S. A. (2009). An encyclopedic dictionary of language testing (4th ed.). Tehran:Tung Hua Book Company
6. Wiggin A. (1993). Testing terms . English Teaching professional. Forum. July. Issue Twenty
7. Forum for Teachers
8. file://A:\Forum\Vol
9. www.ets.org.
10. www.ets.org/TOEFL/

