ASSESSING TEACHERS' TESTING SKILLS IN ELT AND ENHANCING THEIR PROFESSIONAL DEVELOPMENT THROUGH DISTANCE LEARNING ON THE NET

Dinçay Köksal, PhD
Çanakkale Onsekiz Mart University
CANAKKALE-TURKEY

There are so many tests coming from on high that I have no idea what they are. We are doing so many tests and all we find is that the patients are very sick. But what do we do, how do we get the patient well? To do that, we have to have time to teach.

(Cumming 1997, cited in Brindley 2002)

Abstract

This paper discusses what teachers of English do to test their learners with specific test samples used to assess students' language skills, and suggests practical solutions on this issue. It also proposes a distance learning course in 'testing and assessment' for teachers' professional development on the Internet, which will enable them to exchange ideas and find solutions to problems.

Keywords: Distance learners, Distance education, Turkey, English teachers.

Introduction

In the past there was not a course for 'testing and assessment' in FLT departments in pre-service teacher-training programmes. The teacher trainees were given some information about test types but they were not instructed how to design/construct, administer and score the tests when assessing language areas and skills. Within the reconstruction of Faculties of Education in Turkey since 1998, we have had such a course: English Test Construction and Evaluation.

Like me, many graduates of ELT did not know what the characteristics of a good language test were and how to prepare, administer and evaluate language tests. So we, teachers of English, prefered to use ready-made tests constructed by other testers/teachers or the tests offered in the textbooks. The disadvantage of using such tests is that we teach but others test what we teach.

THE PROBLEM: Two Cases

In my 17 year of teaching experience what I have observed is that teachers generally ask their students to perform tasks that extend their limits. Here I would like to report two cases: one that I experienced on a Saturday afternoon in the winter of 2001 when I started to teach the course 'English Test Construction and Evaluation' to the fourth year students, and the other two years later.

Case 1:
I went to my office after teaching a private language course in Foreign Language Education and Research Center and met a woman talking to my secretary about the homework that her daughter was assigned to do. She said that her daughter had to write a composition about *What the world would be like in the year 3000* and added that she had asked a teacher of English, who was her neighbour, to translate what her daughter wrote in Turkish and that the teacher had promised to. But after one week the teacher returned the paper untranslated, stating that the text contained the words and structures she couldn't translate. The mother was therefore forced to ask some else to do this homework. I couldn't help asking what age her daughter was and wanted to see the text the girl wrote in Turkish. The mother said her daughter was a sixth year student aged 11. I had a quick glance at the text and noticed many complex sentences.

No doubt any native speaker at that age can make long and complex sentences and express themselves as effectively and efficiently as an adult. However, the Turkish girl had to use another code to re-express her ideas. It was certain that

the student went beyond her imagination, but it was unfair to expect her to go beyond her limited linguistic capacity.

Case 2:

Due to semantic narrowing, the teachers of English, like many other teachers in Turkey, think that the word 'test' refers to *multiple choice tests.*

I asked the teacher trainees at the beginning of the 2003-2004 academic year to collect some test samples from the schools where they were doing teaching practice. They reported that when they asked their mentors to give them some test samples, the tutors responded "we don't administer tests (refering to multiple choice tests) to our students any longer. We use traditional tests."

These two cases provide enough evidence that teachers of English need training in language testing and evaluation. Such training is crucial because we should regard testing as a bridge between teaching and learning and classroom tests as mirrors in which teachers and students can see their reflections clearly.

Here we should focus our attention on the following basic ***Principles of Testing*:**

In the first case;

*To assess learners' performance in the target language, the teacher should not give a task that the learners cannot perform. The task should be **authentic, realistic** and **appropriate** to their linguistic level.*

In daily life we need to predict future happenings such as weather, exchange rates or interest rates in the coming days or months, economic or political changes in our country or the world, but we do not need to talk about what will happen in the next millenium- the year 3000. The task here was unrealistic.

Even when assessing the learners' performance, at any level, the learners

> *should be given clear instructions well. They should know what they are expected to do in a given task. The ideas, feelings and emotions that the learners want to express cannot be limited to their insufficient linguistic input.*

The mother said that her daughter started to translate what she had written in Turkish but couldn't succeed and asked me to do this homework, adding that if she did not submit it the following day she would fail. It is very traditional in Turkey to ask other people to do work for you and many people do it. I generally refuse to do such things on principle but I wanted to use this case in my course. I agreed to do this task for her daughter and told her to come and collect it back the following morning. Of course, her teacher knew that she couldn't write such a composition using complex expressions and structures and that another teacher of English or some other competent user of English would do it. Unfortunately, the teachers are indirectly assessing their colleagues' performance, not their students'.

> *Teachers should test the outcomes or products of what they have taught their learners, not what their colleagues know*

Data collection

During the academic years 2001-2002 and 2002-2003, 56 different classroom test samples (progress, quiz and achievement) were collected from the state schools where trainees from our ELT department had teaching practice. We observed that the same classroom tests had been used without any revision or editing.

Discussion

Common characteristics of test samples

The tests collected from the schools shared the following points :

1. They did not tell us who they targeted, what skill or area of ability they intended to measure, how much time was allocated; or what points the test-takers would get for each correct response.

2. They did not have separate sections clearly stated, such as 'vocabulary, grammar, writing, reading.' Some had but since the collage techique was used more than two sections had test items for the same skill or knowledge.

3. The test items had more than one possible answer because they were not contextualised.

   For instance, the test-takers were asked to choose between the two:

   I *decided/have decided* to move to a better job.

   I *worked/ have worked* in an office in Japan.

   In these two questions, both options are correct since the test items have no context.

4. The time allocated for each task was not stated on the test papers. We have no idea whether the students had enough time to perform the tasks. Generally only the total time available to perform all the tasks was given.

5. The tests that were hand-written were mostly unreadable. Most were full of mistakes in spelling and grammar. Can test-takers be expected to give a response to something they cannot percieve visually?

6. In the same test, one part contained simple structures while the other contained complex srtructures. *The level of the students was not considered.*

7. The reading texts were too long and included some tasks that did not require any reading skill. For instance, one of them had an exercise that required the test-takers to practise the structure (need+Ving). This had nothing to do with reading comprehension. We know that we can use a reading passage as a starting point for discussion in a conversation class, for dictation, in writing or listening. What was wrong with this test was that the teacher borrowed the reading text with all the exercises

from another text-book without adapting it to the purpose of his test. The section had the title 'Reading' but after the reading text the test-takers were subjected to tasks irrelevant to the skill in question.

8. Most instructions were wrong and unclear.

Another important pitfall we observed in the tests was that the instructions were not given clearly or directly. For instance, in one of the tests we examined the test-takers were asked to 'fill in the blanks' but below that instruction there was nothing to fill in. The task the test-takers needed to perform was to put the words given in the right order. This case proves that teachers do not have time to edit or do not ask other colleagues to proof-read what they have written.

Problems concerning Washback Effect, Validity and Reliability

9. We know that testing has a washback effect on language learning and language teaching. The tasks that we expect our students to perform in classroom activities must be in concert with the tasks they are asked to in tests. In other words, they must be familiar with the tasks and techniques that the teachers use to assess the learners' langauge skills. We should choose tasks in tests by taking into consideration the kinds of instructional activities that the students have been exposed to.

The teacher shouldn't use a technique not used in the teaching process as a test technique to have a positive washback effect of testing on language learning and teaching.

Validity

10. One important principle is that test items must be representative of what we intend to test.

Teachers borrow tests from other books but do not pay attention to the content of those tests. For instance, they teach three types of

conditionals but the test they borrow includes only two types. They administer this test to the students and when they give correct answers to the questions, they conclude that the students have learnt all the conditionals. This test does not, therefore, have content validity.

There were some tests claiming to test the test-takers' writing skills. In fact the task did not require that skill. The test-takers were asked to read the texts, match them with the pictures and then put the numbers in the boxes.

Construct Validity

11. There were some tests claiming to test speaking skills. The test-takers were asked to fill in the missing parts of dialogues. Of course, dialogues are part of our daily communication. However, while we are speaking, we do not present our ideas or feelings on paper. This may be part of a writing task. When I was a first year student in the ELT department, the professor tested our speaking skill by using a paper-pencil test asking us to write dialogues. Twenty years later teachers are still using the same technique.

Teachers should test learners' writing skills by having them write and their speaking skills by having them speak. This is what is known as 'construct validity'.

---

11. There were some tests that looked like a collage. That is what I call the '**collage technique**' in testing. Different test items/exercises were photocopied from different test books and stuck on the paper; and on the top of the first page the name of the school, class, and course written. The reason was that teachers do not have enough time to construct and write tests. It may be a practical way of test preparation, but in fact is impractical when its results are considered.

No matter whether the tests are self-production, collage-production, or publishers' production, it is essential to ask the following questions:

> Is the task perfectly clear?
>
> Is there more than one possible correct response?
>
> Can test-takers arrive at the correct response without having the skill supposedly being tested?
>
> Do test-takers have enough time to perform the task(s)?

How can we rely on results of tests that are unreliable and invalid?

We administer many tests to language learners from the primary level to the secondary, and from the secondary to the higher education and students pass these tests. Finally, learners and teachers are considered to have achieved the goals determined by the Ministry of Education and Higher Education Council, which aim at enabling learners to speak, write and understand what they hear or read in the target language. This is the ideal, but the reality is that there is little or no evidence to show that the goals have been achieved. We still have many graduates of universities or high schools who are expected to use a foreign language but who cannot express themselves even in the simplest way.

Conclusion

As may be understood from the test samples, the reasons for this failure can be listed as follows:

1. Teachers are not trained in testing (test construction, administration and assessment). But one of the most important roles of teachers is as a tester. In private schools there are testing offices and teachers are responsible for teaching only. Achivment tests are constructed, administered and marked by this office. But in many state schools it is the teacher that is responsible for testing and assessing his students. However, they are not able to do this task effectively.

2. Testing and teaching were found not to overlap. We teach something but we test something else. In other words, instructional objectives are not taken into account when choosing test tasks.

3. Tests focus on recognition but not production. Most importantly, our testing focuses on the learning itself, not the outcomes of the learning. As Edge (1996) said,

We teach people and we evaluate language ability but we do not evaluate people.

Suggestions

Now that teachers of English need to be trained in testing and assessment, there is every good reason for the effective use of computers. We are all aware of the important place computers have in almost every aspect of life. In language learning we can use the Internet for communicative practice through student-student interaction (Highashi 1997:78) as well as guided student-computer interaction, which is useful for practising grammar, functions and even lexical items.

The Internet provides an ideal environment for creation and dissemination of interactive language materials (Biddulph, 1997:79). Using the familiar medium of the www, it is possible for foreign language teachers to share their experience and ideas concerning 'testing and assessment'. Here the responsibility must be shared by the department of teacher training, Ministry of Education and Faculties of Education. Both parties can establish a FORUM for discussion of Testing and Assessment on the Internet so that through tester-to-teacher, teacher-to-teacher, teacher-to-teacher educator and teacher-to-student interaction, not only nation-wide but also world-wide collaborative distant learning can be realised.

Of course, teachers can be trained using in-service training seminars and workshops in local or regional in-service training centers. Face-to-face

interaction may be more effective than imaginary interaction on the NET but the advantages of using the NET in teacher training and development are overwhelming:

- It is the fastest way of communicating and not time-consuming.
- Teachers will have many sources of information they can reach whenever they need them.
- It is more practical than in-service training seminars organised during the summer holidays when teachers want to do something else.
- It is the cheapest and easiest way of educating the relevant people (teachers, teacher trainees, teacher educators and testers.

To enable teachers of English to acquire and develop their testing skills so that they can build a strong bridge between learning and teaching, they can be trained through a distance learning course in Testing and Assessment on the Internet. In such a training course I propose there should be two parts:

In the first part, teachers should be made familiar with Action Research as a means of developing reflective classroom practice by introducing key concepts and approaches and enabling them to have an understanding of research in general and develop Action Research skills in the field of Testing and Assessment in ELT. Finally they can write their action research project in the relevant area. As Brindley (1997) states, the involvement of teachers in developing specifications, item writing and trialling can ensure that *test content is consonant with current teaching practices,* thus increasing the likelihood of beneficial washback of the test on teaching.

In the second part, teachers are introduced to the course: aims, prerequisites, course calender, activities involved, useful web addresses, course materials and assignments.

References

Brindley, G. (2002). "Assessment and the Language Teacher: Trends and Transitions." **The Language Teacher Online**.
http://langue.hyper.chubu.ac.jp//tlt9//sep/brindley.html

Bilddulph, P. (1997). "Interactive exercises for language learning." **IATEFL 1997 Brighton Conference Selections.** Edited by Peter Grundy.

Edge, J. (1996). **Essentials of English Language Teaching.** (3rd impression). London: Longman

Genesee, F. and Upshur, J.A. (1998) **Classroom-based Evaluation in Second Language Education.** (3rd edition) Cambridge: C.U.P.

Higashi, A.M. (1997). "Using Internet for communication practice." **IATEFL 1997 Brighton Conference Selections.** Edited by Peter Grundy.

Hughes, A. (1996). **Testing for Language Teachers**. (8th edition). Cambridge: C.U.P.

Weir, C.J. (1990). **Communicative Language Testing.** (2nd edition) New York: Prentice Hall.