
5a. Teaching Educational Testing and Measurement at Tibet University

(Haiwen WANG)

Overview

During the three weeks from June 13 to July 2, 2005, I taught a short course “Educational Testing and Measurement in the English Classroom” at Tibet University. It was designed to introduce college English teachers and students to current and future trends in education and educational testing. Participants were expected to learn how to design tests that yield valid and reliable measurement of how well they teach and how well students learn. Participants were also expected to learn to assess the advantages and limitations of tests and test scores, so that they can interpret test results appropriately and effectively.

The course was designed as consisting of six three-hour sessions. But because English teachers at Tibet University were undergoing a comprehensive teaching review of the academic year, and the students were busy preparing for final exams, the course was condensed to four three-hour sessions, to be delivered to a teacher class and a student class respectively. Due to emergency meetings called by the University, one session was canceled for each of the classes. Therefore three three-hour sessions were actually delivered to the teacher and the student class respectively, totaling 18 hours. My fixed office hours were 10 hours, on top of which I also made myself available during flexible hours for individual attendants.

The teacher class was made up of all those teaching English major and some of the young teachers teaching general English at Tibet University. The total attendance was 20. The lectures were given in English. The student class consisted of 30 first-year students majoring in English Teaching (there were no higher grade students from that major when the course was given). Since these students were not proficient in English, the lectures were delivered mainly in Chinese.

The Course

Since the course was designed to impart an integral concept of testing and measurement, it is inappropriate to discard certain topics due to the time constraint. I therefore decided to teach all the contents despite the reduced teaching time of four sessions. The first session for the teachers was canceled so I further condensed the course to 3 sessions. Unfortunately, the last session for the students was canceled without advanced notice, therefore I did not deliver the last few topics to the student class. Following is a summary of the sessions for the teachers. Those for the students are similar except that theirs were given at a slower pace.

The first half of the first session introduced the attendants to issues in contemporary educational testing and measurement, such as recent trends in classroom measurement, current and future trends in educational measurement and how these trends affect the classroom teacher. The attendants were also informed on the purpose of testing and the importance of testing in the current era of accountability and rising standards. More specifically, the attendants were taught how to set up concrete goals of measurement and what to pay attention to in order to reach those goals.

In the second half of the first session, the attendants were introduced to the concept of norm- and criterion- referenced tests, the different purposes they serve, and how content validity plays a part

in these kinds of tests. The procedures for measuring learning outcomes were also introduced, with an emphasis on the three-stage model of classroom measurement.

The first half of the second session familiarized the attendants with different types of test items. It gave a detailed discussion of the advantages and disadvantages of objective items and essay items, followed by guidelines to write these items, with specific examples. The attendants were also instructed on how to construct rubrics and how to score essay items.

The second half of the second session consisted of a discussion on how to administer, analyze, and improve the test, including assembling, administering, scoring and analyzing the test, together with an overview of the process of evaluating classroom achievement. Marks and marking systems were also introduced, with an emphasis given to the combination and weighting of the components.

The first half of the third session introduced standardized tests and the concepts of variability, normal distribution, and converted scores, together with basic statistical knowledge. Different types of validity and validity coefficients were also introduced.

In the second half of the third session, attendants were introduced to the definition of reliability, methods to estimate reliability, and the interpretation of reliability coefficients. Lastly, the relationship between validity and reliability was discussed.

The planned fourth session for the student class included validity, reliability, and the relationship between them. This session was not delivered due to the cancellation of class by the University.

Teaching Method

The course was originally designed to emphasize student participation. For each session, I prepared a big number of discussion questions and exercises. And participants were expected to form groups to create their own measurement instruments. However, due to the reduced class time, the course was delivered mainly in the form of lecturing to ensure that attendants were familiarized with all the basic concepts. All the prepared reading materials were distributed, with some designated as 'optional'. Most of the attendants made efforts to read all the 'required' materials. A final paper was also required.

Office Hours

The office hours proved to be very helpful to some attendants who were interested in the topic and those who had some practical experiences with testing and measurement. They came in with questions and showed me the instruments they had, which I helped them to analyze and improve. Some made further appointments to discuss educational issues in off-office hours. I personally enjoyed discussing with them and learning what they were exploring in the field of educational testing and measurement.

Feedback from Students

The course was required for teachers who teach English major and the student class. The attendants were not given specific information on the course prior to the sessions. Some attendants even thought the course would discuss teaching methods or classroom management. It took me some efforts to get them into the "testing mood". However, from the attendance rate and in-class and off-class interactions with the attendants, I found that most of them thought the course was useful and made serious efforts to learn.

However, due to the intensity of the course and the discussion of statistical methods in the last session, some attendants found it difficult to follow. It was a pity that I did not have enough time to warm them up on statistical concepts involved or provide them enough opportunity for in-class discussion.

Since the last session for the student class was canceled, I only administered feedback forms to the teacher class. The teachers' feedbacks confirmed what I felt about the course— it was complete and useful in general, but a little too intense, and more elaborations should be made for the attendants to understand the concepts to a better extent.

Thoughts from the experience

The teaching trip to Tibet University was an invaluable experience for me. I am grateful to OYCF for the teaching fellowship to enhance my teaching ability, and more importantly, to bring back to China the knowledge and perspectives I have acquired from the Ph.D. program in the United States. Having been studying and conducting educational research in the US for four years, I found that the US is much more systematic in scientific research in the educational setting than China. In China the general public, parents, students, teachers and educators attach great importance to test scores, but very few of them, not even the teachers or educators, understand what the scores really represent.

To interpret a test score, it is necessary to understand what constructs the test measures, the way it measures the constructs, the validity and reliability of the test, how the test is scored, and the distribution of the scores. Teachers should possess some basic knowledge in testing and measurement before they design, administer, score classroom tests or interpret test results to the general public. Testing is not only a means of evaluating students, but more importantly, it should also be used to guide and inform instruction. It is important for teachers to be capable of designing valid and reliable tests that are sensitive to learning. It is also important for teachers to use test results appropriately and effectively, based on their understanding of the advantages and limitations of tests and test scores. In this sense, knowledge of testing and measurement is vital to today's teacher.

I was prepared to see the attendants not familiar with the concepts in testing and measurement, but was a little surprised to hear a few deem it as 'not relevant to my teaching career'. I think this view came from a lack of emphasis on testing as a means of improving instruction from the policy makers. It is also due to China's centralized management system in education, where teachers do not have much decision-making power. And it is not easy to change from the bottom. On the other hand, I am glad to find that this course did make a difference. Some attendants became very interested in the topics. They started to analyze what they had at hand and think carefully about how they can apply the knowledge in their teaching. The attendants agreed in general that the course was helpful.

The greatest difficulty I had from the teaching trip was the cut of the course from six to three sessions. Although I managed to present an overall picture of testing and measurement, I did not have enough time to further explain the concepts and their applications to the attendants. Besides, there was very little discussion in the sessions. To make up for this problem, I extended my office hours. This was very helpful to some attendants, but not all attendants were able to make full use of it.

Experience with Tibet University

Before I made the trip, I made contact with the university through Wei Hong, the assistant director of the foreign affairs division of Tibet University. She forwarded my course syllabus to the English department and gave me feedback from them (they preferred that I teach both the teachers and students in the department, instead of only students as I original proposed). However, I did not get to talk to the department directly to arrange for the teaching details, and I supposed they arranged the class schedule as I proposed. Instead I finalized the teaching schedule with Professor Xiao in the English department after my arrival.

Wei Hong is a very considerate host. She arranged for me to stay in the foreign student's apartment and had one of her staff show me around. I also met with the dean and other leaders of the English department during my stay. They extended warm welcome to me and showed interest in future OYCF teaching programs. They informed me that the ideal time for the teaching program is May to early June, when schedules are not hectic at the department. The department also invited me to a one-day picnic for the teachers. I came to know my students much better from that event.

As Dr. Hu Xiaojiang of OYCF advised me, students at Tibet University are at a much lower academic level compared with average university students in China. I found it impossible to give lectures in English to the first year students in the department of English. On the other hand, the students are well disciplined. Teaching technology is not advanced in the university, either. I was surprised to find that some students do not have an email account, although all the teacher's offices have fast Internet connection. Tibet University is building a new campus. I expect the facilities to be much better there once it is finished.

(The author is a Ph.D. candidate in the Department of Education in University of California, Los Angeles.)