

## Voices in the Field: An Interview with Lyle Bachman

by Jeff Hubbell

Lyle Bachman is Chair of the Department of Applied Linguistics & TESL at the University of California, Los Angeles. He is on the editorial board of *Language Testing*, one of the most important professional journals in the field, and the author of two books that are essential in any language tester's library: *Fundamental Considerations in Language Testing* (1990) and *Language Testing in Practice* (1996), co-authored with Adrian Palmer. This interview was conducted during his Japan tour in March 2002.



### *In what ways has your interest in testing changed over the years?*

Given that fact that I've been at this business longer than I sometimes care to remember, this is both an interesting and a difficult question, as it causes me to go back a long way. In a nutshell, the ways in which my interests have changed are related to the evolution of my understanding of the issues and the stages through which I've passed in my career. The ways in which my interests have remained unchanged have to do with the nature of the conundrums language testers face.

I was introduced to language testing through two "on-the-job" training experiences, little knowing at the time that these would lead to a lifelong interest. First was the experience as a teaching assistant in a large lecture class at Indiana University. The professor had us read some research on what he called the "cram-exam" syndrome, which mentioned how most students only studied just before an exam. If there is a mid-term and final exam, then most students only study for a couple of weeks per semester. The research suggested that the way to increase the amount of study time, and hence learning, was to give more frequent exams. However, with over 250 students in the class, we could not possibly give more than two essay-type exams. We thus decided to use multiple-choice (MC) exams. The three TAs in the course spent the summer before the course writing enough MC questions based on the course textbooks to create multiple forms of the test. We essentially cut and pasted sections of the textbooks into MC format, put them all up on punch cards then fed them to a computer as an item bank. The professor then told the students that they would have to take an exam on each section of the course at least once, but if they wanted to take an exam more than that, we would give them the best score they received. Furthermore, we would place all old forms of the section exams on reserve in the library! The students couldn't quite believe this. It

*SHIKEN: The JALT Testing & Evaluation SIG Newsletter, 7(2) June 2003. (pp. 12 – 15).*

was like they'd died and gone to testing heaven. At the end of the course, students in the class had taken every section of the test on average about three times, and we ended up giving something like 90% As. This experience instilled in me the importance of linking teaching purpose and testing purpose, and convinced me that this could be done.

The second experience was when I was in Thailand, where Buzz Palmer and I developed a new EAP placement test. What we did was model the existing English tests of the day. To make a long story short, it didn't work out for a variety of reasons. Different parts of the test we developed turned out to be inappropriate for different groups of the intended test takers. The lessons I think I learned from this were first, never trust the current language testing orthodoxies, and second, always begin with the consideration of test use.

My primary and abiding interest in language testing – trying to better understand what makes tests useful for a given purpose – was informed by these two experiences, and has not changed very much over the years. I believe that this will continue to be a central interest for me, since the investigation of test use gets to the very heart of virtually every major issue in language testing: reliability, construct validity, impact, fairness. And it is here where all the difficult conundrums come together. As I've noted elsewhere, these problems have always been complex, and they will persist.

Over the years I have become increasingly interested in the training of new language testing professionals. Thus, in the past decade or so, my interests have focused mostly on mentoring graduate students into the field, and on conducting training workshops for language testing practitioners. In the past, conducting research was my greatest interest, but now, at this particular stage in my career, I'm finding a great deal of satisfaction from working with students and practitioners. This change in focus has come about because of all the misconceptions about language testing and bad testing practices I've seen around the world over the past thirty or so years. It is my belief that the best opportunity we have for improving language testing is through working with the next generation of language testers – both those who will be our successors as language testing professionals, and those who are professionals in other areas, such as language teachers or applied linguists, and who need to use language tests in their work.

***You've taught testing concepts for many years to many people. What testing concept is generally most difficult for novices to grasp?***

Many people would probably expect the answer to this question to be something statistical, like standard deviation, correlation, or hypothesis testing. However, in my experience, the most difficult and important concept in testing is that of validity, in its broadest, post-Messickian sense.

The reason this is so difficult, I think, is that you can't really get a handle on it without facing many complexities and conundrums. Some examples:

1. If we accept the possibility that different test takers respond differently to the same test task, often for different reasons, then does this mean that this task measures different constructs for different test takers?

2. If we change only one or two characteristics of a particular test task, we may find that this dramatically changes the way some test takers respond to it. How does this affect the kinds of inferences we can make on the basis of performance on this task?

3. If different groups of individuals (e.g., men, women, users of different L1s) perform differently on the same test task, what does this say about the inferences we can make? How does this affect the decisions we make on the basis of their performance? How fair will these decisions be?

The level and breadth of knowledge and skills required to understand validity and to be able to assess it as part of testing research and development is not easily learned, and takes dedication, hard work and time. In my introductory language testing course at UCLA, I discuss validity. In virtually every advanced course and seminar I teach, issues of validity and validation get raised and discussed, no matter what the topic is. By the time my students are ready to write their dissertations, they're beginning to have an understanding of validity and validation.

For a language testing professional, an understanding of validity and validation also requires a thorough understanding of the nature of language use and language ability. This, after all, is what we're trying to measure. This deep knowledge of the nature of language, along with an understanding of validity and the knowledge and skills needed to conduct validation research are probably the defining characteristics of a language testing professional. That's why it's so important for us to allocate a significant amount of our time and effort to the training of new language testing professionals. At the same time, we need to keep working with language testing practitioners - professionals in other areas of applied linguistics who need to use language tests in their work - to help them better understand the complexities of language testing at a level that will help them improve their use of language tests.

***Can you identify any new concerns about the way language tests are being administered or test scores are being used these days?***

My guess is that virtually all the ways in which tests can be misused have been discovered. Furthermore, my cynical view is that these misuses are all probably alive and well somewhere in

*SHIKEN: The JALT Testing & Evaluation SIG Newsletter, 7(2) June 2003. (pp. 12 – 15).*

the world today. A recurring, rather than new danger, I believe, is that of "real-life" or "performance" assessment, which is often seen, by new generations of language testing practitioners (and some established professionals, I would add!) as the answer to all our testing problems. I firmly believe that we need to constantly strive to design language testing tasks so as to engage test takers in language use. Nevertheless, I also believe that focusing on tasks alone, to the exclusion of defining the constructs we want to measure, will not lead us to design and develop language tests that are maximally useful for their intended purposes.

***What advice would you give to language teachers striving to narrow the gap between their work in the classroom and the work of language testing specialists?***

It takes a considerable amount of time and effort to acquire the knowledge and skills necessary to be a language testing professional. The question, then, is what level of expertise do language teaching professionals need in order to use language tests ethically? There is really no answer to this question for all times and places; this will depend on the particular language testing situation. In general, however, I would encourage language teachers to read what's available. There are many excellent language testing textbooks that are appropriate for practitioners. In addition, I'd advise them to take advantage of additional training opportunities - workshops and conferences – that may come their way. Finally, don't hesitate to ask a language testing professional for advice. You can locate a language testing professional by contacting the Japan Language Testing Association or the International Language Testing Association.

### **Suggested reading**

Bachman, L. (1990). *Fundamental considerations in language testing*. Oxford University Press.

Bachman, L. (2000). Modern language testing at the turn of the century. *Language Testing*, 17 (1). 1-42.

Bachman, L. & Cohen, A. D. (1998). *Interfaces between second language learning & language testing research*. Cambridge University Press.

Bachman, L. & Palmer, A. (1996). *Language testing in practice*. Oxford University Press.

**HTML:** [http://jalt.org/test/bac\\_hub.htm](http://jalt.org/test/bac_hub.htm) / **PDF:** <http://jalt.org/test/PDF/Bachman.pdf>