

Careers in Language Testing

Alistair Van Moere

Aaron Olaf Batty
 abatty@sfc.keio.ac.jp
 SRB Associate Editor



Alistair Van Moere is the new president of Pearson's Knowledge Technologies Group, and is responsible for the development, delivery, and validation of their automated language tests, including the Versant test (previously known as PhonePass). Prior to his employment at Pearson, Alistair was instrumental in the development of the Kanda English Proficiency Test at Kanda University of International Studies in Chiba, and drew from his experience with its speaking test in his PhD work under Charles Alderson at Lancaster University—work which won him the Jacqueline Ross TOEFL Dissertation Award in 2010.

He spoke with SRB in the last issue about psycholinguistic assessment, and returns in this issue for our new series profiling those who have made a career of language testing. We were pleased that Alistair was once again willing to take some time out of his busy schedule to talk to us.

About 10 years ago you were a lecturer at Kanda University of International Studies in Japan. How did you get from that position to where you are now?

Actually before Kanda University I was at Shane English Schools, first as teacher then Director of Studies. From there I moved to Kanda. Prof Frank Johnson was the Director of the English Language Institute at the time, and he created an environment where teachers could grow as researchers, and resources were made available to worthy research projects. I'm very indebted to Frank for believing in me and providing me the opportunities to succeed. I took over the coordination of the Kanda English Proficiency Test (KEPT), mainly because I was the only person, from among 35 lecturers, who was interested in doing it. Our visiting consultant, William Bonk, inducted me into language testing. I also taught myself a lot of statistics, and prepared for a PhD.

From there I did my PhD at Lancaster University with Charles Alderson, who was an excellent supervisor. I still contact him for advice. Living in Lancaster felt like exile, it's a pretty isolated place! But it's a strong department and I also learned from people like Dianne Wall, Jayanti Banerjee, and my classmate Spiros Papageorgiou. As my PhD was finishing I looked around to work in a testing company, and Ordinate Corporation in California was head and shoulders above everything else. It was a small, technology driven company, and the work going on was just so cool. It was pioneering speech processing technology and it exuded "the future". Soon after I was recruited into the management team at Ordinate we were acquired by Pearson, and so I unwittingly joined the world's largest education company, which has also been very good to me.

You won the 2010 Jacqueline Ross TOEFL Award for best PhD in language assessment. Have you got any advice for researchers undertaking PhDs in language testing?

It's all about getting the right supervisor and your relationship with your supervisor. Make sure you find someone who is responsive, who is going to read your submissions and give you feedback promptly,

and have good discussions with you. You also need to seek out and surround yourself with excellent people. If there aren't any where you are now, then move. I've been very lucky in this respect, at Kanda, Lancaster and now Pearson. For example at the moment Jared Bernstein, our Chief Scientist, is an excellent mentor. He has a wealth of experience, sees straight to the heart of problems I bring to him, and continually challenges my assumptions. Having people like this around you greatly improves the quality of your thinking.

You spoke with us in the last issue about psycholinguistic testing versus communicative testing. How do you anticipate that a testing practitioner—who is responsible for, say, running a placement test for 1,000 students in a university—can reconcile these two approaches?

Teachers have to be aware of the pros and cons of each approach. A test such as the group discussion reflects what happens in the communicative classroom, and provides washback and practice on pair- and group-work. But, it allows students to fall back on personality, practice avoidance strategies, and it's reliable for separating students into no more than two or three bands. It might also disadvantage students that haven't been inducted into that discussion format before, if they are fresh from high school. Also, score gains might be more illusory than we imagine. At Kanda, freshman students typically increased their score by a few points on a score scale of 0-20, but this was largely due to gains on the trait Communicative Effectiveness. They need only be a bit more comfortable with the interaction, and incorporate back-channeling and enthusiasm, to boost their score.

On the other hand, you can test oral proficiency with a series of discrete point items, such as sentence repeats, or reading a passage aloud and then summarizing it, and then scoring the speech for accuracy and fluency. This is a more controlled approach that creates a level playing field on which to evaluate students and allows the examiner to probe proficiency in a standardized way via items of measured difficulty. But, it's less communicative.

I'm not saying that one approach is better than the other. Just that while communicative tests appear more authentic the performances actually mix in a lot of extraneous skills, and the more reliable approach is to control and standardize the assessment. When it comes to a high-stakes assessment, we should take a mixed approach: we want the benefits of high reliability, as well as the benefits of tasks that elicit communicative skills.

You are responsible for the quality of millions of tests being taken around the world, many of which determine people's career or educational opportunities. What keeps you awake at night?

Test crackers. These are the people or agencies who make a concerted effort to take tests, memorize items, and train students to get higher test scores without improving their proficiency. We have plenty of measures in place to counter them, including biometric identity checks, monitoring during testing, and data forensics. We are also investigating whether certain item types which require a rapid, immediate spoken response are more immune to test preparation strategies. But test crackers are nevertheless a threat.

I think people tend to overlook the fact that when it comes to large-scale English proficiency testing, any test score is a combination of two abilities: language proficiency plus test preparation, and test prep is an alarmingly large proportion of the score. Pearson tests are less susceptible to this because we are a relatively new player, compared to the established test providers, and there isn't a test cracking industry built around the Versant tests or PTE Academic. Anyway, the effects of test prep and test designs which counteract test prep is an entirely under-researched area and I'd welcome more attention given to this.

What's your biggest challenge?

Finding exceptional people to hire. I am always on the lookout for test developers who are trained in linguistics and statistics, who have good project management skills, and can work in a business environment.

What do you see happening in the field these days that has you excited?

In language testing, I'm excited by any studies that involve speed or response time. I think this is an undervalued piece of information in language assessment. Two students can get the answer right, but the one who responds twice as quickly may be much more proficient than the other.

In speech processing, there is extremely exciting research in the measurement of soft skills or aspects of the speaker's state of mind. So for example the machine can predict whether the speaker is friendly, likeable, or patient. This is very promising for our clients in the customer-service (call center) industry

Any new developments since last we spoke?

We have just launched a 4-skills, certification business English test for Japan and Korea called E^{Pro}. It is computer-based, automatically scored, and involves interactive item-types such as responding to emails, and providing oral and written summaries. It is just 90-minutes, reports details on subskills, and has clear advantages over other certification English tests currently available. It can be taken in Pearson VUE centers (<http://www.eproexam.com/>).

Thank you for taking the time to speak with us!

