

## WORKSHOP 3

### Quality assurance in test development: a hands-on introduction to task specifications, item writing, trialling and statistical analysis

by

**Gwen Caudwell:** Aptis Product Development Manager, British Council

**Kevin Rutherford:** Aptis Test Production Manager, British Council

**Judith Fairbairn:** Test Development Consultant, British Council

**John Tucker:** Aptis Test Production Co-ordinator

**Number of participants: min 10 and max 30**

#### Intended learning outcomes:

This practical workshop will help participants learn some of the important quality assurance techniques that professional test developers employ when designing tasks and specifications, writing items, trialling tests, and conducting statistical analysis.

#### Content:

A hands-on workshop where the Aptis team will take participants through the steps to create one reading test.

- Part 1 (3 May): Task specifications
- Part 2 (4 May): Item writing
- Part 3 (4 May): Trialling
- Part 4 (5 May): Statistical analysis

#### Methods:

- An expert in each area from the Aptis team will deliver each part (we will pay for the extra people to attend).
- Participants will be placed into five groups with six people in each group.
- Aptis presenters will circulate and sit with groups, assisting where needed.
- All activities will be done on-line using SurveyMonkey and Google docs (or similar online technology).
- Part 1: Designing and developing a reading task and task specifications.
  - Brief introduction to designing and developing a task and task specifications. Participants will be given a handout on how to write good specifications and a specifications template.
  - Each group will develop one task and write specifications for the task. The task must be no longer than 15 minutes to answer.
- Part 2: Item writing.
  - Brief introduction to item writing. Participants will be given a handout on how to write good quality items and use on-line analysis tools.
  - The task specifications created in Part 1 are passed to a different group.
  - Groups will write items in SurveyMonkey (or similar) using the task specifications.

- Groups will share their items with the group that wrote the specifications and discuss any issues.
- Part 3: Trialling a task
  - Brief introduction to trialling.
  - Groups take the three tasks they have not worked on, responding to the task items. The full test will have been sent the night before to a pre-selected group of test takers in order to collect enough data (or dummy data will be created for the statistical analysis).
  - All marks are automatically collated into a shared template for analysis in Part 4.
  - Discussion to include any problems with items and tasks.
- Part 4: Statistical analysis techniques for trial results.
  - Brief introduction to statistical analysis of receptive skills.
  - Steps to analyse data
    - Descriptive statistics: participants will use excel to calculate descriptive statistics and histograms. Step-by-step handout provided and participants can practice steps.
    - How to create item labels, person labels and a concurrent worksheet in excel for Winsteps. Step-by-step handout provided and participants can practice steps.
    - Create a Winsteps Control File, Item Map and Item Measures. Demonstration only.
    - Discussion to include issues with the data analysis and the test created by the group.

### **Background knowledge/pre-workshop activities**

No specific knowledge required but participants must have a laptop with internet and be setup with Google Docs for sharing work.

**Pre-workshop activities:** None required.

### **Workshop facilitators:**



**Judith Fairbairn:** Test Development Consultant, British Council

She has an extensive ten years of testing experience in managing examiners for IELTS and Aptis. Her areas of expertise include examiner management global policy creation and marking quality assurance, and she holds a certificate in Theory and Practice in Language Testing from Roehampton University.

She is currently studying an MA in Language Testing focusing on examiner marking quality assurance and creating on-line examiner support, training and standardisation systems. Judith was invited to the International Association of Applied Linguistics 2014 conference to deliver a presentation on Quality Assurance in online, international Large Scale Testing. In addition to this, she is also an experienced examiner, item writer, and has provided editorial assistance for IELTS Research Reports Volume 12.



**Gwen Caudwell:** Aptis Product Development Manager, British Council

Ms. Gwendydd Caudwell is Aptis Product Development Manager for the British Council based in Dubai. In 2009 she took the Roehampton Theory and Practice of Language Testing course as a means to understand the rationale behind some of the tests she was using with her students in her teaching career.

This sparked her interest in this area and she completed her MA in Language Testing with Lancaster University in 2012. Gwendydd has extensive experience in teaching, teacher training and management for the British Council around the world.

She is responsible for the research and design of new variants of Aptis and British Council tests and has also worked on designing and implementing tests in different contexts such as the Science Olympiad Foundation in India and a Young Learner test for a large project in Uruguay. Her particular areas of interest are rater judgements and Young Learner assessment and is looking to base her PhD on this topic.



**Kevin Rutherford:** Aptis Test Production Manager, British Council

Kevin Rutherford is currently the Aptis Test Production Manager, and is based in Warsaw, Poland. He is responsible for managing the production of content for Aptis tests and for the trialling of test items, and the setting up of Aptis test centres.

He also leads on the training of item writers and is involved in the development of new test items. Before moving into the field of assessment, he had a long career in English language teaching, in Japan, Italy, and the United Kingdom, and, since 1996, in Poland.

He completed the Certificate in the Theory and Practice of Language Testing with the University of Roehampton in 2010, and was subsequently involved as an item writer in the early development of Aptis. He obtained an MA in Language Testing from the University of Lancaster in 2014. He has given presentations on assessment at conferences for the International Association of Teachers of English as a Foreign Language (IATEFL) and the European Association for Language Testing and Assessment (EALTA).

In 2014 he co-presented a workshop on item writing and test specifications at the Language Testing Research Colloquium (LTRC).



John Tucker has been Aptis Test Production Co-ordinator, based in Poland, since July 2012. His current responsibilities include commissioning and quality reviewing test items for Aptis General and other tests, along with helping to produce the tests themselves. John is involved in producing materials for and training item writers, and presented at ALTE 2014 in Paris on this subject. He also handles technical queries connected with the administration of Aptis tests. John is part of the British

Council's new Assessment Literacy Project, for which he is writing a module on listening; John also co-presented on this topic at a pre-conference TEASIG workshop in Harrogate. Earlier in 2014 he took an open access course in corpus linguistics run by Lancaster University.

John completed an MA (distance) in Language Testing as well as the BC's E-moderator Essentials course in December 2013. Before that he took a diploma in testing from Roehampton University. Previously, John ran placement testing workshops and helped re-design assessments for the British Council's Korea centre in Seoul. Prior to that, John had a long and varied career as a teacher in the UK, Colombia, Spain and Thailand.

**CVs are on the following website:**

<http://www.britishcouncil.org/exam/aptis/research/assessment-group>