# Research Notes

## Contents

## Editorial Notes

Welcome to issue 38 of *Research Notes*, Cambridge ESOL's quarterly research publication. This issue focuses on professional development, both within Cambridge ESOL and more widely. This topic includes the services we offer to encourage and support the professional development of educators and other stakeholders working with those taking a Cambridge ESOL examination or teaching award. More generally, this theme encompasses teacher education.

In the opening article, Juliet Wilson, Cambridge ESOL's Assistant Director of Customer Services, describes how the Customer Services teams train and support staff at our worldwide network of centres to ensure that all candidates receive a consistently high quality experience when taking any of our exams or teaching awards.

The next three articles consider teacher education. Firstly, Richard Rossner discusses the rationale for developing common criteria when comparing the aims and content of teacher training programmes for in-service teachers and prospective teachers of foreign languages, making reference to the EAQUALS profiling grid for language teachers. Next, Christine Coombe, Mashael Al-Hamly and Salah Troudi focus on teachers' knowledge of assessment principles which is commonly referred to as *assessment literacy*. They discuss some of the challenges and offer recommendations for ways forward in this area. Laura Mureşan then provides a case study of teacher development within the context of higher education in Romania, focusing on teachers of English for Specific Purposes or English for Academic Purposes.

We then consider Cambridge ESOL's outreach services and provision to stakeholders worldwide. Andrew Nye and Karen Barns describe the support that we offers our teaching community; they report the results of a Teacher Survey which gathered views on current professional development services and sought opinions on further developments. Next, Sacha DeVelle describes the procedures followed when selecting materials for writing and speaking examiner training, standardisation and certification, highlighting the mixed method approach followed in compiling these materials.

Next, we focus on studies that provide evidence for our high quality language testing products. Lucy Chambers describes the process behind developing assessment criteria for BULATS (the Business Language Testing Service) which attempts to link criteria to descriptors from the Common European Framework of Reference (CEFR). Laura Cope then examines the reliability of the writing and speaking components of the Computer Based BULATS test.

We finish this issue with conference reports and details of Cambridge ESOL PhD scholarship winners.

Editorial team for Issue 38: Fiona Barker, Hanan Khalifa and Caroline Warren.

# Support and training for Cambridge ESOL exam centres

**JULIET WILSON** CUSTOMER SERVICES, CAMBRIDGE ESOL

## Introduction

At Cambridge ESOL, a lot of effort and activity is directed towards the professional training and standardisation for examiners of speaking and writing tests and towards support for teachers involved in preparing candidates for our exams, see articles by Nye and Barns on our teacher support and professional development provision, DeVelle for the training and certification of IELTS examiners, and Chambers and Cope's articles on developing assessment criteria and establishing the reliability of a specific test of business language. This edition of *Research Notes* bears witness to much of this work across a variety of contexts including contributions from external authors on the nature of foreign and second language teachers' assessment literacy (Coombe, Al-Hamly and Troudi), developing teachers in a higher education context in Romania (Mureşan) and developing assessment criteria for language teacher education in a European context (see Rossner's article). However, there are other extremely important stakeholders in the examination process whose support and training is not so often reported on. These people are the Centre Exams Managers, supervisors, invigilators and inspectors who work at exam centres. Getting the training and standardisation right for these people is critically important. A candidate's experience on the day of the exam may well be the main impression of Cambridge ESOL that they carry away with them. Moreover, ensuring that we facilitate a bias for best for candidates and optimum exam day conditions for them is no mean feat. This article outlines how the Customer Services teams at Cambridge ESOL train and support our centre staff to ensure that candidates across the world receive a consistently high quality exam day experience.

## Cambridge ESOL Centre Network

Every year more than three million candidates take our exams at our network of 2,500 centres in 130 countries around the globe. Our centre network is very diverse – we work with very large 'open' centres who administer the exams to hundreds of preparation centres, through to small, internal centres based in universities or colleges of higher education, who offer the exams to their students only. We work with private language schools and state-funded institutes, in workplaces and in businesses. Each of these centres plays a crucial role in delivering our examinations and we value all of their contributions. For the Customer Services teams, the challenge is to offer appropriate support and training for each centre.

## Centre Support

The Customer Services teams at Cambridge ESOL are set up to ensure that centres receive excellent ongoing support before, during and after the exam day. Figure 1 shows the structure of our customer-focused team which includes:

- skilled Helpdesk staff who are the first point of contact for all centre queries

- the specialist Application Support team who provide help and advice regarding Cambridge ESOL software systems and computer-based tests and an on-call service on computer-based exam days

- Centre Support Officers deal with more complex enquiries and complaints and offer a personal account management service for centres. Centre Support Officers also design and run the centre training and communications programmes and write the comprehensive documentation, such as Handbooks and Examination Instructions Booklets which detail the strict quality procedures which centres must comply with.

Centres can also count on help and support from our regional offices and development managers around the world.

**Figure 1: Structure of Cambridge ESOL's Customer Services teams**
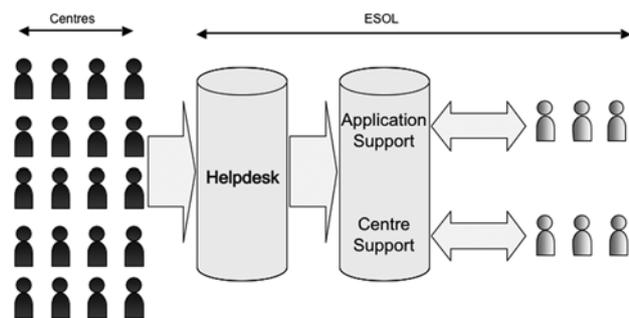


The ESOL Helpdesk uses HEAT, a call logging system, to record details of all emails and phone calls coming into the Helpdesk. This data is an invaluable source of information, allowing us to highlight the areas which centres ask us about most frequently so that these can be included in forthcoming training sessions or clarified in centre documentation. Figure 2 shows the range and number of enquiries dealt with over a one month period. HEAT reports also show the Customer Services teams which centres are calling most frequently and therefore which Centre Exams Managers might benefit from an invitation to attend a training event or to complete online training modules. Cambridge ESOL publishes Service Level Agreements related to response time for queries and complaints so that

**Figure 2: Areas of enquiries to the ESOL Helpdesk during September 2009**
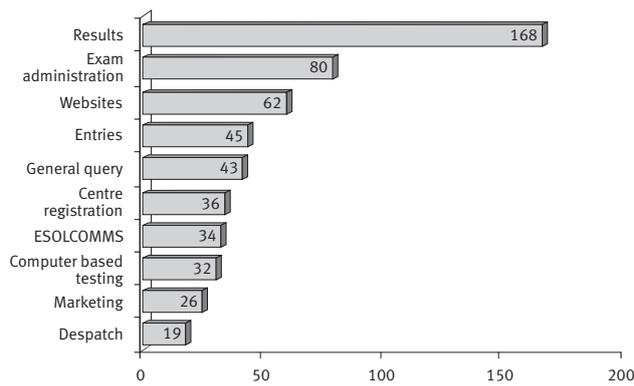


centres are fully informed about the service and support they can expect.

## Centre Training

Running exams in a secure, fair and consistent way is probably the most important part of centre operations. To ensure that centres are fully prepared for the exam day, we offer a comprehensive programme of training. Face-to-face training is offered in the UK on several dates throughout the year and in-country training is also carried out wherever possible. The training sessions are run over two days and are designed to give Centre Exams Managers – particularly those new to the job – a full understanding of the exam administration cycle, from recruiting examiners to making exam entries and processing enquiries on results. In addition, sessions on marketing the exams, briefings about new exams and services as well as a tour of our warehouse facilities – where exam papers are printed, candidates' responses stored and where documentation despatched from – are included in the programme. The content of the training is continually updated by taking into account feedback from previous sessions. For example, a session on best practice in offering exams for candidates with special requirements is now included in the standard training agenda.

For those centres wishing to run computer-based tests, training is held both for test administrators and technical support staff, including hands-on practice in installing the software and running a 'dummy' test. This ensures that centres are fully competent and confident on the day of the exam and can offer an excellent service to their candidates.

Feedback from these training events is always positive, with Centre Exams Managers commenting on the amount of learning which takes place:

'Overall very interesting and gave a clear understanding of how Cambridge ESOL operates. Having the opportunity to speak to experienced exams managers was very helpful, but also to have others who were new to the position meant that we could discuss how we were approaching the new challenges.'

'A full and interesting programme that clarified a great deal. Good opportunities for meeting people both professionally and socially.'

Of course, attending regular face-to-face training may not be possible for every Centre Exams Manager. We are now

exploiting new technology to ensure that exams administration staff can have on-demand access to training and support. We currently use Moodle to deliver training for computer-based tests but over the next few months will be migrating our training modules to Fronter, the Learning Management System which is currently used for training and standardisation of our Oral Examiners.

Online modules for Centre Exams Managers will include 'Your responsibilities', 'Preparing your venues and rooms', 'Staff training' (to be used for supervisors and invigilators) and for Inspectors: 'Preparing for an Inspection', 'During an Inspection', 'Filling in your claim forms' and 'Case Studies'. The training modules use interactive videos with worksheets, graphics and quizzes to ensure that they are dynamic and eye-catching as well as getting the important messages across clearly and in a memorable way. A screenshot from Fronter is shown in Figure 3.

**Figure 3: A screenshot from Fronter**



## 2010: Making Life Easier

Supporting those who work in our centre network is an important priority for Cambridge ESOL and this is borne out in the investment we are making in the *2010: Making Life Easier* project. This project aims to revolutionise the way Cambridge ESOL and exam centres work together – and one of the key stated objectives is to make centres' lives easier. The first phase of the project is the introduction in 2010 of Cambridge ESOL Online – a new website where centres will be able to access all of the tools and information they need to administer the exams. This will reduce the amount of paperwork for both centres and Cambridge ESOL. For example, centres will now be able to create 'confirmation of entries' with personalised details which can be emailed directly to candidates rather than be sent as a hard copy from the UK. Cambridge ESOL Online has been designed in close consultation with centres, which means that its design and features reflect how Centre Exams Managers work. Rather than imposing a new way of working, we are looking to support and enhance centres' current administration procedures. In order to support the transition from our previous exam administration system, we are using the latest technology for training, including WebEx sessions (online and interactive training sessions),

online tutorials via Flash videos and downloadable quick guides. The second phase of this project includes a website for preparation centres, which will further reduce the administration burden for centres, allowing schools to make entries and receive their results directly.

## Continuous communication and improvement

The Customer Services teams are not only involved in providing support and training to Centre Exams Managers, examiners, invigilators and inspectors, but are also committed to using the information we gain to inform new ideas on how we can improve our services. Some recent ideas are to provide online marketing materials for centres and preparation centres, as well as an initiative to introduce online forms to reduce the need for printing and sending paper versions. The regular communications we have with centres means that we are always attentive to what is happening within and outside our centre network. We have recently introduced a dedicated microsite to complement the monthly *Bulletin for Centres*, see Figure 4. This microsite details the latest developments of 2010: Making Life Easier and a short, monthly poll, which will help us gather feedback on all of our new initiatives.

**Figure 4: Screenshot of Bulletin for Centres microsite**



## Conclusion

Cambridge ESOL recognises the indispensable work that staff at centres carry out. As the exams provider, we are aware of our responsibility to offer excellent support and training opportunities for these stakeholders. By doing this, we can ensure not only that we are promoting an ethos of good customer service, which centres can then pass on to their candidates, but also that the quality of the exams we offer is complemented by the quality of exam day processes.

# Developing common criteria for comparison and assessment in language teacher education

RICHARD ROSSNER THE EUROPEAN ASSOCIATION FOR QUALITY LANGUAGE SERVICES, EAQUALS

## Introduction

This article outlines the rationale for a framework for making it easier to compare the aims and content of teacher training programmes for teachers and would-be teachers of foreign languages, and to generate valid criteria for assessment and self-assessment for such teachers. It also briefly describes some background work and points of reference for such a framework. In particular, it refers to the *EAQUALS Profiling Grid for Language Teachers*, and other EAQUALS work in this field. In doing so, emphasis is placed on the fact that any framework for language teacher education needs to take into account existing general frameworks and approaches to the description of teacher knowledge and competencies, and existing teaching standards developed with general and mainstream education in mind.

## The EAQUALS Profiling Grid for Language Teachers

The European Association for Quality Language Services, EAQUALS, is an international not-for-profit association of providers of language courses. Founded in 1991, it now has a membership of over 100 accredited providers across 22 countries and 20 Associate Members – international bodies such as the British Council, Goethe-Institut, and Instituto Cervantes, ten national associations of language schools, and examination boards such as Cambridge ESOL.

The primary aim of EAQUALS is to accredit and support high quality in language teaching and learning, irrespective of the languages being taught and the type of institution. In order to underpin its accreditation of language teaching institutions, EAQUALS has developed a very thorough and comprehensive system of quality assessment through on-site inspection and review of documentation. The *Inspection Scheme Manual*, now in its sixth version, provides detailed criteria for the reference of schools and inspectors, as well as a checklist of focus points and sources of evidence. These criteria are based on the EAQUALS Charters, four documents that are displayed by Accredited Members, and which spell out in simple terms the standards that the institution promises to maintain in its general conduct and values, in its services to course participants, in its relations with staff members, and in the promotional and other information it provides. A school inspection normally takes two days and involves two

**Table 1: Basic layout of the *EAQUALS Profiling Grid***

| Level of development | Basic | | Independent | | Proficient | |
|---|---|---|---|---|---|---|
| | T1 | T2 | T3 | T4 | T5 | T6 |
| **Language**<br>• **Language proficiency**<br>• **Language awareness** | | | | | | |
| **Qualifications and experience**<br>• **Teacher training qualifications**<br>• **Teaching practice**<br>• **Experience** | | | | | | |
| **Core competencies**<br>• **Methodology – knowledge and skills**<br>• **Lesson and course planning**<br>• **Interaction management and monitoring**<br>• **Assessment** | | | | | | |
| **Complementary skills**<br>• **Teacher development***<br>• **Use of digital media***<br>[*for example – various other possible skills are suggested] | | | | | | |

inspectors from different countries. Accredited Members of EAQUALS are re-inspected every three years.

As part of its commitment to providing support for its members and would-be members, EAQUALS has developed a comprehensive set of training and reference materials, covering in particular such topics as: applying the principles of the Common European Framework of Reference (CEFR; Council of Europe 2001) in developing curricula and means of assessment for language courses; introducing and using the *European Language Portfolio* that EAQUALS developed in collaboration with the Association of Language Testers in Europe (ALTE); and staff development and appraisal. In 2004 it was decided that, to facilitate standardised quality assessment across the Association irrespective of the national context, the diverse background of teachers, and the target language, a simple means of describing teachers' qualifications and competences at different levels was needed. Brian North, a co-author of the CEFR and current EAQUALS Chair, and Galya Mateva, an experienced teacher trainer and Chair of the Bulgarian Association for Quality Language Schools, OPTIMA, decided that the best means of doing this would be to develop a grid of the kind used in the CEFR. They then produced a first draft of the Profiling Grid, which was presented at the EAQUALS International Conference in Athens (November 2005).

The rationale for and genesis of the EAQUALS Profiling

Grid is described in detail in North 2009. As indicated in Table 1, its horizontal axis deliberately follows the pattern of the CEFR self-assessment grid in encompassing three main levels of development, while its vertical axis covers four broad categories: 'language', 'qualifications and experience', 'core competencies', and 'complementary skills'.

As can be seen from the whole *Grid*, version 30 (2009) of which is reproduced in Table 2, the cells in the *Grid* contain descriptors of the competencies, training, experience etc. for each of six levels, where T1 and T2 are seen as 'training' levels, and T5 and T6 are levels where substantial experience has been accumulated and greater teaching competence has been attained. As with the CEFR and language learners, it is not expected that the profile of a given teacher will be uniform: it is far more likely that she (or he) will be stronger in some areas than others. Typically, a native speaker of English, for example, will have proficiency in the language which is at a higher level than her 'awareness' of it from the point of view of learning and teaching. Similarly, a given teacher may be more competent in the area of planning than she is in interaction management, or vice versa. Thus, it is assumed that the profile of most teachers will be 'jagged'.

The main use of the Grid has been to enable managers to provide profiles of teachers in their team prior to an inspection. The Profiling Grid is issued with a request that institutions to be inspected describe their teachers' qualifications and competencies with reference to it, thus enabling inspectors to have a more comprehensible and relevant overview of the teaching team than would be possible if they had to rely on reading through everyone's CV. This is important as the aim of observations and other evidence gathered about teachers during an EAQUALS inspection is to form an overview of the team's ability to maintain the standards laid out in the charters and in the inspection criteria. It is not to identify individual strengths and weaknesses (indeed, inspectors' oral and written reports provide no feedback on individual members of staff). If they wish, schools can produce a scatter-gram or summary as indicated in Table 3 to provide a composite assessment of their teaching team.

Thus, the main aims and benefits of the EAQUALS Profiling Grid for Language Teachers have been:

• to provide easily interpretable information about the teaching team at an institution for inspection purposes

• to enable comparison of teacher qualifications within countries and internationally

**Table 3: Sample of possible means of using the *Profiling Grid* to describe a team of teachers prior to EAQUALS inspection**

| | | Basic | | Independent | | Proficient | |
|---|---|---|---|---|---|---|---|
| | | T1 | T2 | T3 | T4 | T5 | T6 |
| **Language** | **Language Proficiency** | | Teaching assistant 1 | Teachers F, H | Teachers B, D, G, K | Teachers A, C, J | Teacher E |
| | **Language Awareness** | | | Teacher B, H T. Assistant 1 | Teachers F, C, D, K | Teacher G, J, E | Teacher A |
| **Qualifications** | **Language Teacher Qualifications** | | ✔ | ✔ ✔ ✔ | ✔ ✔ ✔ ✔ ✔ | ✔ ✔ | |
| | **Language Teaching Practice** | | ✔ ✔ | ✔ ✔ ✔ ✔ | ✔ ✔ ✔ | ✔ | |
| | **Teaching Experience** | ✔ | | ✔ | ✔ ✔ ✔ ✔ | ✔ ✔ ✔ | ✔ |

- like the CEFR itself, to provide a common language and shared points of reference in terms and levels.

Presentations and workshops outside EAQUALS meetings in several countries and for different kinds of audience have given rise to considerable interest in the concepts behind and potential of descriptive tools of this kind. In particular, the innovative conferences on Training, Quality and Certification in Foreign Language Teaching (TQAC in FLT) run by the University for Foreign Students, Siena, in 2008 and 2009, which brought together teacher trainers working in

**Table 2: Draft EAQUALS Profiling Grid for Language Teachers (Brian North, Galya Mateva, Richard Rossner 2007–2009, Draft 0.30)**

| | | Basic | | Independent |
|---|---|---|---|---|
| | | **T1** | **T2** | **T3** |
| **Language** | **Language Proficiency** | • studying the language at tertiary level<br>• B1 proficiency | • studying the language at tertiary level<br>• B2 proficiency | B2 certificate in the language; oral competence at C1 level |
| | **Language Awareness** | • answer simple queries with the help of reference works | • answer queries related to high frequency structures | • give correct models of usage on most occasions<br>• answer most language queries satisfactorily at A1–B1, using reference sources as necessary |
| **Qualifications** | **Language Teacher Qualifications** | • taking a certificate in teaching the target language,<br>*or:*<br>• following an internal training course | a minimum of 30 hours documented, structured training in language awareness and methodology of teaching the target language | a minimum of 60 hours of documented, structured training in teaching the target language |
| | **Language Teaching Practice** | experience of team-teaching or of acting as a teacher's assistant | experience of supervision and assessment while teaching phases of lessons | • a minimum of 2 hours of documented, assessed teaching practice<br>• has been observed and had feedback on some actual teaching |
| | **Teaching Experience** | taught some lessons or parts of lessons at one or two levels | own class(es) but limited experience which only includes teaching at lower levels | • a minimum of 200 hours, documented teaching experience<br>• taught a range of levels up to B1 |
| **Core competencies** | **Methodology: knowledge and skills** | • sensitisation to learning theories and features of language<br>• familiarity with a limited range of techniques and materials for one or two levels | • basic understanding of learning theories and features of language<br>• familiarity with techniques and materials for 2+ levels<br>• select new techniques and materials with advice from colleagues | • familiarity with theories of language learning and with learning styles<br>• familiarity with an expanding range of techniques and materials<br>• choose which to apply based on the needs of a particular group<br>• evaluate usefulness of techniques and materials in teaching context |
| | **Lesson and Course Planning** | work with lesson plans in teachers' notes to published materials | • use published or in-house materials to develop plans for different types of lessons<br>• plan phases and timing of various lesson types | • use a syllabus and specified materials to prepare lesson plans that are well-balanced and meet the needs of the group<br>• adjust these plans as required<br>• take account of lesson outcomes in planning next lesson |
| | **Interaction Management and Monitoring** | alternate between whole class teaching and pair practice following suggestions in a teachers' guide | • manage teacher-class interaction effectively<br>• give clear instructions for pair and group work<br>• monitor the resulting activity<br>• give clear feedback | • set up pairs and groups efficiently<br>• ensure all learners are involved in productive pair and group work<br>• monitor performance at all times<br>• bring the class back together and manage feedback |
| | **Assessment** | supervise and mark class quizzes and progress tests | • supervise and mark tests<br>• write a class quiz or revision activity to revise recent work | • select suitable progress tests and set up and supervise them<br>• use the results and simple oral and written tasks to assess learners' progress and things to work on<br>• use a homework marking code to increase language awareness |
| **Complementary skills** | **Teacher Development** | • take part in training sessions<br>• cooperate with colleagues with set tasks<br>• regularly observe real teaching | • take an active part in group work during training<br>• liaise well with other teachers<br>• observe and team-teach with teachers at restricted levels<br>• act on observation feedback | • take an active part in various kinds of in-service training/development<br>• actively seek advice from colleagues and relevant books<br>• observe colleagues at various levels<br>• act on colleagues' feedback on serial observations of own teaching |
| | **Digital Media** | • write a worksheet foll-owing conventions<br>• follow menus to operate software<br>• download from resource sites (e.g. One Stop English) | • search effectively for material on the internet<br>• select and download from resource sites (e.g. One Stop English)<br>• organize materials in hierarchical folders | • use data projectors for class lessons with internet, DVD etc.<br>• use software for handling images, DVDs, sound files<br>• use a camcorder to record tasks<br>• set a class an exercise with CALL materials |

several languages, provided opportunities for reflecting on the commonly felt need for a coherent set of criteria, and possible ways forward, as well as for comparing current assessment practices. Simultaneously, within EAQUALS the desire for a means of describing and comparing language teacher training courses internationally has grown.

## Descriptive tools for language teacher training – the precedents

Overlapping with the genesis of the EAQUALS Profiling Grid, two other projects have produced comprehensive instruments that aim to lay out the territory for language

| | Proficient | |
|---|---|---|
| **T4** | **T5** | **T6** |
| a C2 examination certificate (e.g. CPE) | • degree in the language, *or:* <br>• a C2 examination certificate (e.g. CPE) | • native speaker, *or:* <br>• language degree or C2 certificate plus a natural command of the language |
| • give correct models of usage on most occasions <br>• answer language queries adequately though not always comprehensively, using reference sources as necessary | • give correct examples of usage on all occasions <br>• answer language queries reliably | • provide clear explanations <br>• teach usage and register at all levels <br>• understand what is confusing learners <br>• give comprehensive, accurate answers to queries |
| • degree in the target language, *or:* <br>• internationally recognised (minimum 100 hours) certificate in teaching the target language | • degree or degree module in teaching the target language, *or:* <br>• internationally recognised (minimum 100 hours) certificate in teaching the target language | • masters degree or module in language teaching or applied linguistics *or:* <br>• postgraduate or professional diploma in teaching the language (minimum 200 hours) |
| • a minimum of 6 hours of documented, assessed teaching practice <br>• has been observed and had feedback on at least 5 hours of real teaching | • a minimum of 12 hours of documented, assessed teaching practice <br>• has been observed and had feedback on at least 8 hours of teaching | • a minimum of 18 hours of documented, assessed teaching practice <br>• has been observed and had feedback on at least 12 hours of teaching |
| • a minimum of 800 hours, documented teaching experience <br>• taught all levels except C1 and C2 | • a minimum of 2,400 hours, documented teaching experience <br>• taught all levels except C2, exam, and/or specialised classes | • a minimum of 4,000 hours, documented teaching experience <br>• taught all levels successfully, general, exam and specialised |
| • familiarity with learning theory, learning styles and learning strategies <br>• identify the theoretical rationale behind a wide range of techniques and materials, with which familiar <br>• evaluate appropriateness of techniques and materials in different teaching situations | • good familiarity with teaching approaches, learning styles, strategies <br>• provide theoretical rationale for teaching approach and for a very wide range of techniques/materials <br>• evaluate materials effectively from practical and theoretical perspectives | • detailed knowledge of theories of language and learning <br>• select an optimum combination of techniques to suit each type of learner and learning situation and provide clear theoretical rationale for decisions |
| • analyse individual learners' needs in detail, including learning-to-learn <br>• plan clear main and supplementary objectives for lessons <br>• provide a rationale for lesson stages <br>• select/design supplementary activities <br>• ensure lesson-to-lesson coherence | • plan a balanced, varied scheme of work for a module based on detailed needs analysis <br>• design tasks to exploit linguistic and communicative potential of materials <br>• design multi-level tasks to meet individual needs and lesson objectives | • plan an entire course with recycling and revision <br>• create or select appropriate activities for balanced learning modules with communicative and linguistic content <br>• design multi-level tasks to meet individual needs and lesson objectives |
| • set up a varied and balanced sequence of class, group and pair work appropriate to the lesson objectives <br>• monitor individual and group work effectively providing or eliciting appropriate feedback | • set up group interaction focused on multiple learning objectives <br>• monitor individual and group performances accurately and thoroughly <br>• give various forms of relevant individual feedback | • facilitate task-based learning <br>• manage learner-centred, multi-level group work <br>• derive appropriate action points from monitoring and analysis of the interaction |
| • conduct tests and interviews if given material to do so <br>• train learners to code their errors to increase language awareness <br>• design or select appropriate quizzes, revision activities, and progress tests <br>• CEFR standardisation experience | • coordinate placement testing and progress assessment (oral and written) <br>• use video and homework codes to help learners recognise strengths/weaknesses <br>• use CEFR criteria reliably to assess spoken and written proficiency | • write progress tests <br>• develop assessment tasks <br>• run CEFR standardisation sessions <br>• use video and homework codes to help learners recognise strengths/weaknesses <br>• use CEFR criteria reliably to assess spoken and written proficiency |
| • develop awareness and competence through professional reading <br>• lead discussions sometimes and exchange ideas about materials and techniques <br>• seek opportunities to be observed and receive feedback on own teaching | • act as mentor to less experienced colleagues <br>• lead a training session or even series of sessions given materials to use and distance support from a colleague <br>• seek opportunities for peer-observation | • create a series of training modules for less experienced teachers <br>• run a teacher CPD programme <br>• take part in institutional or (inter) national projects <br>• observe colleagues and provide effective feedback |
| • use data projectors for class lessons with internet, DVD etc. <br>• create lessons with downloaded texts, pictures, graphics, etc. <br>• set and supervise individual CALL work <br>• coordinate project work with media (camcorder, intent downloads etc.) | • use PowerPoint for presentations, including animation <br>• train students to select and use CALL exercises effectively <br>• use authoring program to create CALL <br>• troubleshoot with basic equipment (e.g. data projector, printer) | • show colleagues how to use new soft/hardware, including authoring programs <br>• design blended learning modules <br>• use any standard Windows software, including media, video editing <br>• troubleshoot hardware |

teacher training in different contexts. Both refer specifically to initial language teacher training provided by universities. *The European Profile for Language Teacher Education – a Frame of Reference* (EPLTE) was the result of an EU-funded project:

'It deals with the initial and in-service education of foreign language teachers in primary, secondary and adult learning contexts and it offers a frame of reference for language education policy makers and language teacher educators in Europe. The findings draw on consultation with a wide range of European experts on language teacher education, and on the experience of eleven European teacher education institutions. The findings also suggest guidelines for quality assurance and enhancement. By outlining the key elements in European language teacher education, the Profile aims to serve as a checklist for existing teacher education programmes and a guideline for those still being developed.' (Kelly et al. 2004:4)

EPLTE contains 40 criteria or 'indicators' grouped into four sections:

- Structure: elements of a training programme, and modes of training and learning.
- Knowledge and Understanding: background knowledge in terms of language proficiency, knowledge about language and learning, technical issues such as information and communication technology (ICT), curricula, evaluation.
- Strategies and Skills: practical knowhow in terms of methodology, etc.
- Values: intercultural, collaborative, and European values that teachers should be trained to promote in their teaching.

The body of EPLTE is a series of descriptors related to 'should' statements aimed at policy-makers, and followed by rationales and examples from case studies collected from among the project partners. For example, under Knowledge and Understanding, we find:

'Foreign language teacher education in the twenty-first century should include the following elements of initial and in-service education:

14. Training in language teaching methodologies, and in state-of-the art classroom techniques and activities

Explanation
- Trainee teachers learn about and employ different language teaching methodologies.
- They know the different ways of achieving learning outcomes, and the different techniques necessary for teaching reading, writing, speaking and listening and for improving reception, production, interaction and mediation skills in learners.
- Trainee teachers learn how to use up to date classroom techniques and activities based on interactive, group, and peer-assisted learning.

Elaboration
- Trainee teachers who learn about a number of methodological approaches to teaching and learning are able to adapt to particular contexts, and have a firm foundation for the critical and creative use of teaching theories.
- Trainee teachers also learn about different methodologies and new classroom techniques from peers training to teach different languages, where methodologies and resources vary.' (Kelly et al. 2004:46).

The same pattern is followed for all headings.

By contrast, the *European Portfolio for Student Teachers of Languages* (EPOSTL) is described as:

'a document for students undergoing initial teacher education. It will encourage you to reflect on your didactic knowledge and skills necessary to teach languages, helps you to assess your own didactic

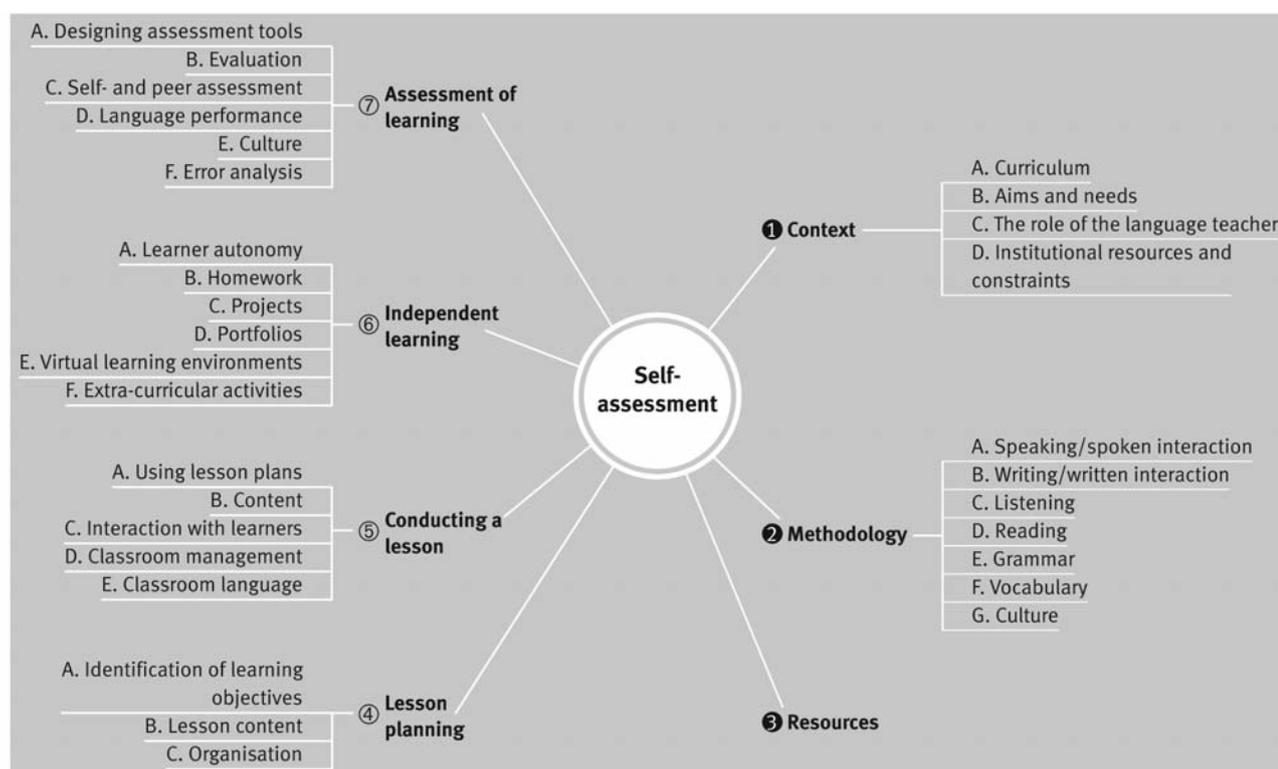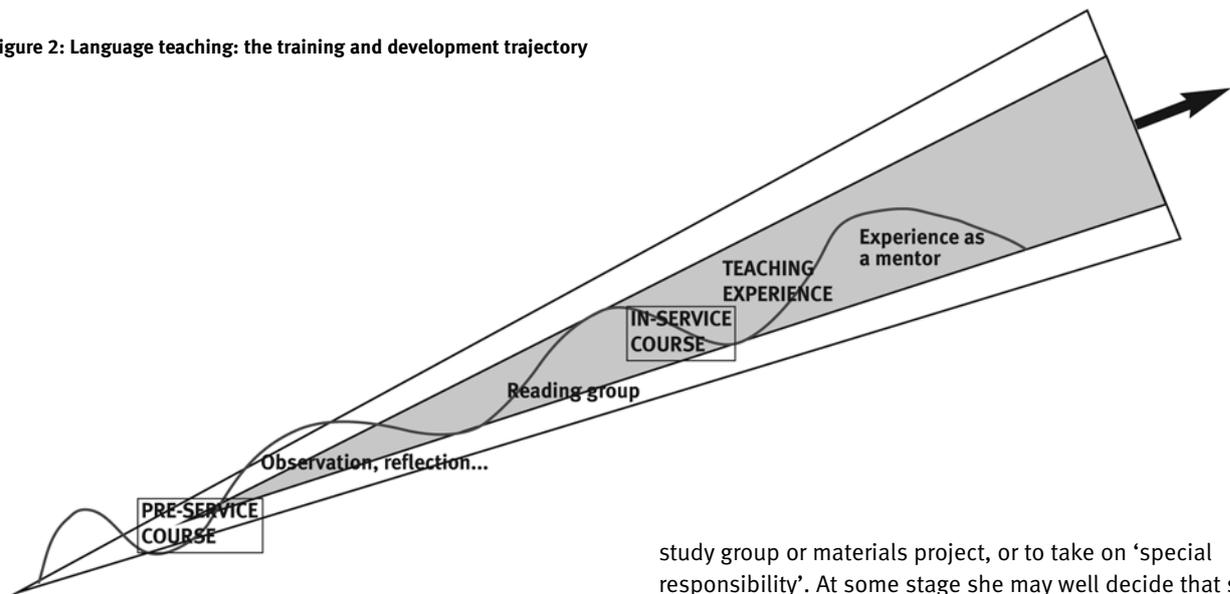**Figure 1: Outline of EPOSTL descriptors (EPOSTL 2007:6)**

**Figure 2: Language teaching: the training and development trajectory**



competences and enables you to monitor your progress and to record your experiences of teaching during the course of your teacher education.' (Newby et al. 2007:5)

EPOSTL resulted from a project run under the auspices of the Council of Europe's European Centre for Modern Languages (ECML) in Graz. Mirroring the approach used in European Language Portfolios (ELPs), a Council of Europe-inspired initiative to enable students of languages to asses their own progress using descriptors based on those in the CEFR, the focus of EPOSTL is clearly on self-assessment by student teachers, who are invited to use EPOSTL to track their own progress against 193 descriptors: 'These descriptors may be regarded as a set of core competences which language teachers should strive to attain' (Newby et al. 2007:5). The 'map' of the descriptors is outlined in Figure 1.

## Towards an EAQUALS Framework

The professional career of a language teacher (or any teacher) can be depicted as indicated in Figure 2. Within the context of her (or his) rich and busy life, she takes a decision to go into the teaching profession to teach a second or foreign language, and (usually) takes a course or courses in preparation for this, either as part of a university degree or as a professional qualification. During or shortly after this training, her experience as a language teacher begins, represented in Figure 2 by the shaded triangle. It may well start in a 'narrow' way – teaching at one or two levels within a single institution. If the teacher in question is lucky, there will be good induction support and ongoing mentoring from a director of studies, head of department, or an experienced teacher, who will observe some lessons, provide helpful feedback, offer guidance with materials selection and so on. As time goes by, the teacher will gain wider experience by teaching classes at new levels, in different age groups, or with specific needs, and she may well move from one institution to another.

As indicated by the wavy line in Figure 2, our hypothetical (and perhaps idealised) teacher would expect to have opportunities to participate in in-service training workshops, to carry out peer observation, take part in a

study group or materials project, or to take on 'special responsibility'. At some stage she may well decide that she wishes to broaden and deepen her knowledge and skills as a language teacher, and to enrol for a course at Diploma or Masters level. This in turn gives her new development experiences and horizons. Later on in her career the teacher may take on responsibilities beyond teaching, involving, for example, the mentoring of less experienced teachers, materials or test creation, co-ordination of courses or academic management.

In the one-page *EAQUALS Profiling Grid* it was not feasible to cover all the areas encompassed by EPLTE and EPOSTL. However, unlike these, the *Grid* attempts to demonstrate, and to some extent define, a progression from less experienced and less proficient to very experienced and proficient, thus making the tool relevant to teachers at all stages of the development trajectory depicted in Figure 2, and their trainers. But the *Grid* remains primarily a tool for providing basic 'standardised' information about teachers working in a given organisation.

Considering the potential of the concept of international criteria and descriptors for language teacher training and development, an EAQUALS Project Group[1] has done further work with the aim of eventually developing tools that will:

a. Facilitate comparison between different teacher training programmes for teachers of different languages working in differing contexts, both at pre-service and in-service level, and provide a firmer basis for the various means of and criteria for assessment used on such courses.

b. Assist institutions in the development and description of their training programmes for language teachers, in the way that language courses are now planned and described with reference to the CEFR.

c. Provide criteria against which such teacher training programmes could be quality-assured, in the context of an independent and international accreditation scheme.

d. Expand on the EPOSTL concept by providing tools for language teacher self-assessment in the context of their continuing professional development (CPD).

---

1  The author wishes to acknowledge the contributions of the members of the EAQUALS Special Interest Group who have been working on this project, especially their ideas and draft descriptors, some of which are referred to below. The project group will be meeting again in November 2009, and work will continue in 2010.

e. *(linked to d)* Provide employers with sounder means of 'appraising' teachers and working with their employees to plan and provide appropriate personalised CPD.

Aims a, b and c, above, echo to some extent a recommendation in the EPLTE Final Report:

'a European-level evaluation framework for initial and in-service teacher education programmes, enabling accreditation and mobility ... . A framework [should be] established allowing programmes, courses and modules to be evaluated at a European level. ... Its role will be to ensure recognition and transferability of teacher education qualifications throughout Europe.' (Kelly et al. 2004:38)

Because of EAQUALS' goal to support language teachers at all levels, any EAQUALS framework would need to address knowledge and skills in more detail and more broadly than either EPLTE or EPOSTL did, and include teachers being trained to teach, or already teaching, their own mother tongue (explicitly excluded from EPLTE). The EAQUALS Project Group began by compared the *EAQUALS Grid* with both EPLTE and EPOSTL for content and coverage. Several areas were identified that were not included in the *Grid* but would need to be covered in the proposed framework to make it sufficiently comprehensive. On the other hand, it was found that, in general, both EPLTE and EPOSTL were less explicit and practical in their focus than the Grid where core classroom competencies are concerned. It was provisionally decided that, unlike the existing *EAQUALS Profiling Grid*, the proposed EAQUALS framework should include the following:

- Subcategories of descriptors derived from subdividing the Grid's 'methodology – knowledge and skills' and 'interaction management and monitoring' sections into subcomponents. Subdivisions might include, for example:
  - theory behind different approaches and methods (i.e. pedagogic 'knowledge')
  - competence in applying the principles of the CEFR in classroom teaching
  - methods for teaching vocabulary, grammar, and pronunciation (language knowledge) as opposed to language skills
  - learner training, including 'learning how to learn'
  - teaching monolingual classes and multilingual groups
  - teaching young learners, university students, company employees etc.

Of these, some work has already been done on the second and fourth.

- Additional 'complementary skills', i.e. skills that relate to the wider context of teaching and learning, some of which were already mentioned in the notes on the original *Grid*. These could include, for example:
  - learner counselling
  - quality management
  - language testing
  - materials development
  - managing people, e.g. as a co-ordinator.

- Areas not dealt with at all in the *Grid* but referred to in EPLTE and/or EPOSTL, for example:
  - interpersonal skills, such as building a good rapport with students, and being a good listener

- intercultural awareness, including facilitating students' introduction to the target culture, appreciating cultural diversity, and encouraging students to do so, and willingness to learn about the cultural background of students
- teacher professionalism, including willingness to develop and evaluate yourself, openness to feedback from students and management, and contribution to the school and wider community.

Some initial work has been done on these.

The Project Group also concluded that it would be important to look beyond language teaching at the various taxonomies and standards produced for teachers in general, bearing in mind that a growing majority of language teachers internationally work in the context of mainstream education, from kindergarten and primary level upwards[2]. The Group is thus currently looking at four sets of standards, two drawn from the UK, one from the US, and one from Australia, described below.

### Professional Standards for Teachers (PST), UK Teacher Development Agency (2007)

The introduction to the above Standards states that:

'The framework of professional standards for teachers ... defines the characteristics of teachers at each career stage ... Professional standards are statements of a teacher's professional attributes, professional knowledge and understanding, and professional skills ... The standards provide the framework for a teacher's career and clarify what progression looks like ... To access each career stage a teacher will need to demonstrate that he/she has met the relevant standards.'

Like the EAQUALS *Grid*, but on the much larger canvas of primary and secondary education in England and Wales, PST, then, sets out cumulatively what is expected at different levels (five in this case) of experience, expertise and career progression as a public-sector teacher. Like EPLTE, which has sections on Knowledge and Understanding, Strategies and Skills, and Values, PST divides the standards as follows:

a. Professional attributes
b. Professional knowledge and understanding
c. Professional skills.

Here is an example of standards for 'professional skills' at the Core (second) Level:

'Teach challenging, well-organised lessons and sequences of lessons across the age and ability range they teach in which they:

a. use an appropriate range of teaching strategies and resources, including e-learning, which meet learners' needs and take practical account of diversity and promote equality and inclusion

b. build on the prior knowledge and attainment of those they teach in order that learners meet learning objectives and make sustained progress

c. develop concepts and processes which enable learners to apply new knowledge, understanding and skills

d. adapt their language to suit the learners they teach, introducing new ideas and concepts clearly, and using explanations, questions, discussions and plenaries effectively

---

2  It is not clear from the documentation whether the preparatory work for EPLTE or EPOSTL reviewed such systems and standards, as no references to any are listed.

e. manage the learning of individuals, groups and whole classes effectively, modifying their teaching appropriately to suit the stage of the lesson and the needs of the learners.' (PST 2007:19)

This single standard (out of 41) for 'Core Teacher' is both demanding and, from an EAQUALS point of view, thought-provoking. Everything here can apply to language teachers working in different contexts, and each implies 'core competencies' that need to be addressed in teacher training and teaching practice. But questions also arise: what is a 'challenging' lesson? What kind of evaluation of teaching materials and resources is implied by '...take practical account of diversity and promote inclusion'? And what sort of training does the development of these professional skills entail?

### New Overarching Professional Standards for Teachers, Tutors and Trainers in the Lifelong Learning Sector (PSTLL) – Lifelong Learning UK (2007)

This related but differently focused and framed set of standards aims to provide a basis for qualifications and career progression in the very diverse lifelong learning and further education sector in the UK. The expectations of and qualifications for such teachers were, to say the least, ill-defined, and levels of competence, and effectiveness were also disparate. Here there are no different levels of attainment related to career progression. However, the standards are divided across seven domains covering what teachers might be expected to do within the lifelong learning sector, although it is conceded that not all teaching jobs in the sector will involve doing all that is implied here:

- Domain A: Professional values and practice
- Domain B: Learning and teaching
- Domain C: Specialist learning and teaching
- Domain D: Planning for learning
- Domain E: Assessment for learning
- Domain F: Access and progression.

'Professional values and practice' are seen as underpinning all the other domains, each of which is also divided into three sections. At the head is 'professional values' referring back to domain A, and below these in two linked columns are 'professional knowledge and understanding' and 'professional practice'. This is interestingly different from the example from PST above (aimed at teachers in mainstream schools), particularly in its hierarchical organisation: domain B is linked back to the 'overarching domain' through the five values spelt out in the first section, which, as in domain A, focus on learners, learning, equality, reflection and collaboration; these are then unpacked and elaborated on from the point of view of 'learning and teaching' as professional knowledge and professional practice are described in more detail. Nevertheless, professional practice is described at a fairly general level (even more general than in PST) when considered from the standpoint of a framework for teacher training:

'BP 2.1 Provide learning activities which meet curriculum requirements and the needs of all learners.

BP 2.2 Use a range of effective and appropriate teaching and learning techniques to engage and motivate learners and encourage independence.

BP 2.3 Implement learning activities which develop the skills and approaches of all learners and promote learner autonomy.

BP 2.4 Apply flexible and varied delivery methods as appropriate to teaching and learning practice.' (PSTLL 2007:6)

What training is implied by 'Implement learning activities which develop the skills and approaches of all learners and promote learner autonomy', and, equally importantly, what criteria would be used to evaluate whether or not a teacher has met this part of the standard? These are important questions when considering the interpretation and use of such standards. However, from an EAQUALS Framework point of view the standards provided here and in PST are useful checklists against which to identify gaps in criteria specifically aimed at teachers of languages, and stimulate debate about how to plug them.

### The National Board for Professional Teaching Standards (NBPTS), USA

This independent entity has developed a large array of standards for teachers in the US covering different specialist areas as part of a certification system for teachers. According to the website, 'National Board Certification: meets most states' definition of "highly qualified teacher" under NCLB [The No Child Left Behind Act 2001]; strengthens teaching practice; improves students' learning according to a vast majority of research; advances teaching careers; [and] increases financial opportunities in many states and districts'. In other words, there are strong reasons why teachers may well be keen to seek certification under these standards.

Unlike the UK standards outlined above, NBPTS standards are presented by subject specialism ('certificate areas', of which there are 25), implying necessary differences between standards for different kinds of teachers. The Standards are each presented in booklet form. The one exemplified here, *English as a New Language for 3–18 Year Olds*, runs to 146 pages, including space for 'reflection' on each standard. For example, Standard 5 is:

'Meaningful Learning: Accomplished teachers of linguistically and culturally diverse learners use a variety of approaches that allow students to confront, explore, and understand important and challenging concepts, topics, and issues in meaningful ways.'[3]

The Standard is followed by three pages of continuous prose elaborating and exemplifying it (and two blank pages for reflection). For example:

'Teachers use their past experience along with their knowledge of children and various subjects to develop an ongoing mix of activities, discussions, and social interactions that allow children to create their own understanding of what they learn. They model the kind of creative thinking and problem solving that will enable children to become successful in their own endeavors. They are skilled at observing; listening; facilitating discussion; orchestrating play; asking questions;

---

3  This is similar to standard 5 in the booklet for generalists: 'Accomplished generalists require students to confront, explore, and understand important and challenging concepts, topics, and issues and to improve skills in purposeful ways', and presumably to the equivalent standard in all certificate areas, although there are differences in the text that follows.

adapting materials and routines to new use; and helping children make connections with past ideas, experiences, and bodies of knowledge.' (NBPTS 2001:25)

The considerable space used for elaborating on each standard is helpful in enabling teachers (and readers looking for input into standards of their own) to understand what might be implied by a given standard. However, the actual means of assessing whether an individual teacher meets the standard are not so clear. Nevertheless, there is much here that would provide a useful check on elements in an eventual framework for foreign language teacher training.
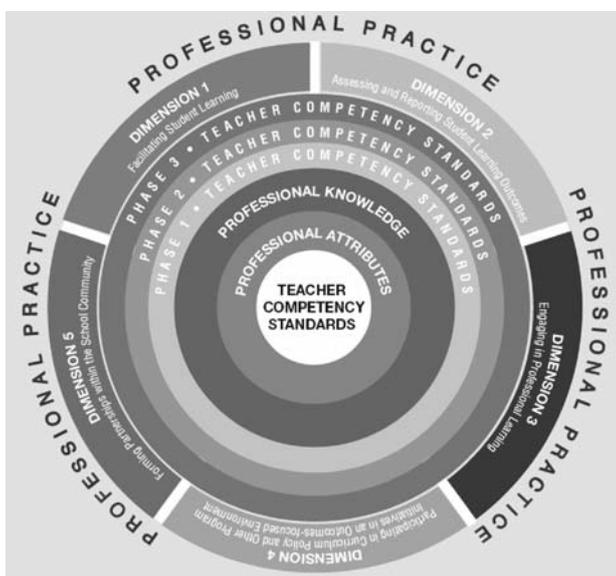
### Competency Framework for Teachers (CFT), Department of Education and Training, Western Australia (2004)

The CFT is at first sight less focused on providing a framework for performance management career progression. The introduction states: 'the Department's development of this Framework [CFT] provides teachers, teacher educators, teacher organisations and professional associations with a description that establishes agreed dimensions of effective teaching and offers a common reference point for professional reflection, discussion and action' (CFT 2004:iii). The standards set out in CFT follow on from the much less detailed *National Framework for Professional Standards for Teaching* produced by the Australian Government (2003). They are laid out in three 'phases' of development in a teacher's career with five different 'dimensions' at each phase. The framework is colourfully illustrated in Figure 3, and the first three dimensions (the other two do not relate directly to the classroom) are laid out in Table 4. Nevertheless, they are occasionally referred to as 'standards' and, as Table 4 shows, the word 'should' is used frequently.

A gradually more detailed approach is used to present the competencies in CFT. For example, at Phase 1, 'facilitating student learning' is defined as:

'Phase 1 teachers facilitate student learning by planning lessons that engage students and provide a purpose for learning. They experiment with different approaches to teaching, addressing the needs of students and priorities of the school. In this phase, learning is often teacher

**Figure 3: Competency Framework for Teachers – Western Australia**



**Table 4: Summary of 'professional practice' from the Competency Framework for Teachers (2004:8)**



| | PHASE 1 Teachers operating within the first phase should: | PHASE 2 Teachers operating within the second phase should: | PHASE 3 Teachers operating within the third phase should: |
|---|---|---|---|
| **DIMENSION 1** Facilitating Student Learning | Engage students in purposeful and appropriate learning experiences | Cater for diverse student learning styles and needs through consistent application of a wide range of teaching strategies | Use exemplary teaching strategies and techniques that meet the needs of individual students, groups and/or classes of students in a highly responsive and inclusive manner |
| **DIMENSION 2** Assessing and Reporting Student Learning Outcomes | Monitor, assess, record and report student learning outcomes | Apply comprehensive systems of assessment and reporting in relation to student attainment of learning outcomes | Consistently use exemplary assessment and reporting strategies that are highly responsive and inclusive |
| **DIMENSION 3** Engaging in Professional Learning | Reflect critically on professional experiences in order to enhance professional effectiveness | Contribute to the development of a learning community | Engage in a variety of learning activities that promote critical self reflection and the development of a learning community |
| **DIMENSION 4** Participating in Curriculum Policy and Other Program Initiatives in an Outcomes-focused Environment | Participate in curriculum policy and program teamwork | Provide support for curriculum policy or other program teams | Provide leadership in the school by assuming a key role in school development processes including curriculum planning and policy formulation |
| **DIMENSION 5** Forming Partnerships within the School Community | Establish partnerships with students, colleagues, parents and other caregivers | Support student learning through partnerships and teamwork with members of the school community | Facilitate teamwork within the school community |

directed with the teacher taking responsibility for determining what students will learn, to what degree and how.' (CFT 2004:9)

This is then presented in more detail:

'A teacher operating within this phase is able to:
- undertake planning to support student learning
- apply a professional knowledge base to the design of learning experiences
- promote student learning
- cater for individual student learning styles and needs
- manage teaching and learning processes
- select and use instructional resources and information and communication technologies (ICT).' (CFT 2004:15)

Finally this is broken down into six sub-areas, and 'indicators of effective practice' are provided, for example:

'UNDERTAKE PLANNING TO SUPPORT STUDENT LEARNING
**SOME INDICATORS OF EFFECTIVE PRACTICE**

A teacher operating within this phase:
- identifies learning outcomes for individuals and groups that are matched to students' developmental needs and are consistent with system, district and school curriculum requirements
- prepares purposeful and sequential learning experiences that integrate learning areas and are responsive to student interests and learning styles
- establishes coherent links between intended learning outcomes, learning experiences and forms of assessment
- makes provisions inclusive of individual students with particular learning needs
- addresses student safety issues and concerns.

PROMOTE STUDENT LEARNING

**SOME INDICATORS OF EFFECTIVE PRACTICE**

A teacher operating within this phase:

- makes the purpose of learning experiences explicit to students and links new concepts to prior knowledge
- employs effective questioning strategies to promote student involvement and critical thinking
- uses varying patterns of interaction within and across learning experiences
- emphasizes language as a vehicle for learning through the use of discussion and by listening and responding to the ideas of others
- offers clear explanations of concepts, relationships, procedures and processes
- implements learning experiences that provide opportunities for student collaboration, problem solving, inquiry and creativity
- encourages students to take increasing responsibility for attaining learning goals
- assumes different roles in the instructional process based on the content and purposes of learning experiences
- promotes students' awareness of their own thought processes and the use of reflection to build new understandings.

MANAGE TEACHING AND LEARNING PROCESSES

**SOME INDICATORS OF EFFECTIVE PRACTICE**

A teacher operating within this phase:

- establishes and communicates explicit expectations for student learning and provides clear directions and instructions to students to support their engagement with learning experiences
- structures learning experiences to ensure students have a sense of purpose
- establishes and maintains a classroom environment which has clear, consistent expectations for standards of behaviour
- organises, allocates and manages time, materials and physical space to support learning
- encourages students to assume responsibility for their behaviour
- uses praise and encouragement to foster students' self-esteem and to promote investment in effort
- facilitates student ownership of classroom procedures
- applies positive management techniques to respond to off-task or inappropriate behaviour'. (CFT 2004:16–7)

In this respect, CFT is closest to what EAQUALS aimed at with the Profiling Grid for Language Teachers, and it contains features that are very useful reference points in putting together a coherent framework for language teacher training and development.

## Validation, assessment, accreditation

It is evident from these four documents for mainstream education and from EPLTE, that defining the scope and content of any framework of descriptors or standards for language teacher training will be a considerable undertaking. However, there is evidence in all these examples that the task is feasible and worthwhile.

In EAQUALS' case, key issues remain to be considered.

How can such criteria be validated, and to what extent should language teachers themselves (or indeed language learners) be involved in the process of validation?[4] From an EAQUALS' perspective, given the diversity of language teachers and teaching contexts represented, it would be essential to undertake a structured consultation exercise with a cross-section of such teachers. Another potential aid to validation would be the numerous resource books that

exist for teachers undertaking training or doing self-help development. For example, many exist for teachers of English to speakers of other languages, including, for example, *The Practice of English Language Teaching* (Harmer 2007), *Tasks for Language Teachers* (Parrott 1993), and *A Course in Language Teaching* (Ur 1999). However, in referring to them, the specific culture of English language teaching as projected by books of this kind would need to be taken carefully into account.

It is clear from the examples referred to that the standards are intended to play a key role in assessment, particularly assessment related to career progression, as in the case of PST or NBPTS[5]. In the case of a framework for teacher training, the criteria and descriptors provided would need to be suitable both as input for curriculum design and as a basis for certain kinds of course-related assessment and certification, bearing in mind that the validity of atomistic (as opposed to holistic) assessment of teaching is likely to be questioned, not least by teachers themselves.

A third area referred to in EPLTE and already under consideration is how these standards fit together with criteria that could be used for accrediting teacher training courses, which brings to the fore other issues such as the structure and balance of the course, teacher training methodology and modes of training, the organisation and supervision of teaching practice, and the way trainees are assessed.

## Conclusion

The EAQUALS Profiling Grid was a landmark in the evolution of EAQUALS' policy on the description of foreign language teacher competencies, and is useful in preparing for school inspections. It is now having a significant impact on EAQUALS' thinking about language teacher training and development, and ways in which the effort to raise the quality and effectiveness of such training and development can be supported.

A descriptive framework of reference to meet EAQUALS members' and partners' potential objectives needs to be:

- comprehensive in its coverage of the competencies (or skills), knowledge (or awareness) and values (or attitudes) required of teachers of languages, including relevant general teaching competencies
- applicable to – or capable of expansion to – various educational sectors, ranging from pre-primary to adult learning, and from specialised language courses for professional purposes or content-and-language-integrated learning (CLIL) to intensive general courses
- capable of encompassing no fewer than three levels of experience and proficiency as a teacher, beginning with a 'basic' or trainee level
- intuitively valid, and endorsed by teaching professionals,

---

4  Interestingly, NBPTS states that the members of committees appointed to develop and review its standards are in the majority classroom teachers. Similarly, the introduction to CFT states that 'From the outset, the Department acknowledged that in order for the Framework to be a credible and valuable tool teachers needed to have a significant role in the development process'.

5  In the latter case, the website provides interesting insights into the kind of assessment procedures used, which include a portfolio of classroom practice, and online assessment.

as well as respectful of differing teaching and learning cultures

- useful as a tool for:
  - describing and comparing different language teacher training programmes, both pre-service and in-service
  - designing new specialised or general teacher training courses of both categories
  - self-assessment by teachers of their own evolving competencies, knowledge and values
  - informing teacher performance management and CPD in an institutional context
  - creating criteria and modes of assessment for teachers undertaking or ending training
  - generating teaching standards for language teachers in specific contexts where required
  - supporting quality assurance and quality development in teacher education
  - combining with criteria for accreditation of teacher training courses.

The EAQUALS Project Group is aware that this is a considerable undertaking, and not one to be taken lightly. Resources, especially in terms of time and expertise, as well as concentrated effort, will be needed, and consultation and collaboration with partners outside the association is also important. However, success in this enterprise will not just build on the good work done by the EPLTE, EPOSTL and Profiling Grid teams but could provide a very valuable 'open system' resource for the profession in support of continuous improvement in the effectiveness of language teaching and by extension, language learning.

## References and further reading

Council of Europe (2001) *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*, Cambridge: Cambridge University Press.

Department of Education and Skills, W. Australia (2004) *A Competency Framework for Teachers*, available online http://policies.det.wa.edu.au/Members/mullas/policy.2006-01-24.0767983365/Orig_2006-01-24.3228193233.pdf

Harmer, J (2007) *The Practice of English Language Teaching* (4th Ed.), Harlow: Pearson Longman.

Kelly M, Grenfell, M, Allan, R, Kriza, C and McEvoy, W (2004) *European Profile for Language Teacher Education – A Frame of Reference* final report, available online http://ec.europa.eu/education/policies/lang/doc/profilebroch_en.pdf

Lifelong Learning UK (2007) *New Overarching Professional Standards for Teachers, Tutors and Trainers in the Lifelong Learning Sector*, available online www.lluk.org/documents/professional_standards_for_itts_020107.pdf

Ministerial Council on Education, Employment Training and Youth Affairs, Australia (2003) *A National Framework for Professional Standards for Teaching* available online www.mceecdya.edu.au/verve/_resources/national_framework_file.pdf

National Board for Professional Teaching Standards (2001) *English as a New Language –Standards*, available online www.nbpts.org/userfiles/File/eaya_enl_standards.pdf

Newby, D, Allan, R, Fenner, A, Jones, B, Komorowska, H and Soghikyan, K (2007) *The European Portfolio for Student Teachers of Languages*, EPOSTL, available online www.ecml.at/mtp2/FTE/pdf/STPExtract.pdf

North, B (2009) *A profiling grid for language teachers* – presentation to the International Meeting on Training, Quality and Certification in Foreign Language Teaching, Siena 2009, available online http://clients.squareeye.com/uploads/eaquals/North-%20TQAC.pdf

North, B and Mateva, G (2005) *A Profiling Grid for Language Teaching Professionals*, EAQUALS.

Parrott, M (1993) *Tasks for Language Teachers – A resource Book for Training and Development*, Cambridge: Cambridge University Press.

Training and Development Agency for Schools (2007) *Professional Standards for Teachers*, available online www.tda.gov.uk/teachers/professionalstandards/downloads.aspx

Ur, P (1999) *A course in Language Teaching – Practice and Theory*, Cambridge: Cambridge University Press.

# Foreign and second language teacher assessment literacy: Issues, challenges and recommendations

**CHRISTINE COOMBE** DUBAI MEN'S COLLEGE, UNITED ARAB EMIRATES

**MASHAEL AL-HAMLY** KUWAIT UNIVERSITY, KUWAIT

**SALAH TROUDI** GRADUATE SCHOOL OF EDUCATION, UNIVERSITY OF EXETER, UNITED KINGDOM

## Introduction

It has long been acknowledged that assessment is an integral part of the teaching-learning process (James, McInnis and Devlin 2002). In fact, Cowan (1998) calls assessment the engine that drives learning. One of the effective ways of enhancing learning within higher education is through the improvement of assessment procedures.

Research shows that the typical teacher can spend as much as a third of their professional time involved in assessment or assessment-related activities (Cheng 2001, Herman and Dorr-Bremme 1982, Stiggins and Conklin

1992). Almost all do so without the benefit of having learned the principles of sound assessment (Stiggins 2007).

Now more than ever our educational systems are under pressure to be accountable for student performance and to produce measurable results. Without a higher level of teacher assessment literacy, we will be unable to help students attain higher levels of academic achievement. In this paper, we address some issues and challenges related to assessment literacy.

## Current stakeholder views of language assessment

### How students view assessment

For many students, assessment is not an educational experience in itself, but a process of guessing what the teacher wants (McLaughlin and Simpson 2004). For the typical EFL/ESL student, assessment is generally seen as something done to them by their teachers. Many students see tests as threats to their competence and as something to be 'got through'. The more able students enjoy the experience but most students, no matter what their level, feel anxious and worried about assessments as there is great pressure in today's educational world to succeed. When tests or assessments are high-stakes, students often suffer from high levels of test anxiety.

### How teachers view assessment

Teachers often experience similar feelings to those of their students. For those teachers who are not involved in setting tests or assessments for their students, they feel that a gap between teaching and testing is in evidence. They often feel that those who write the tests are not in touch with the realities of the classroom. Research by Jacobs and Chase (1992) found that testing and assessment-related activities are the least fun area of their job.

### How educational boards view assessment

Virtually every set of standards of teacher competence developed recently, including those developed by the National Education Association (NEA), the American Federation of Teachers (AFT), the Council of Chief State School Officers (CCSSO), the National Council for Accreditation of Teacher Education (NCATE), and the National Board of Professional Teacher Standards (NBPTS) have identified and endorsed a set of assessment competencies for teachers (Wise 1996 as cited in Stiggins 1999).

In the field of English language teaching, TESOL partnered with the National Council for Accreditation of Teacher Education (NCATE) created the TESOL/NCATE Standards for ESOL teacher education. Assessment constitutes one of the five knowledge domains within these standards. In Europe, the Common European Framework of Reference and the European Portfolio for Modern Languages are requiring language teachers to adopt new ways of assessing language ability (Stoynoff and Coombe forthcoming).

Clearly, there is widespread global recognition that language assessment literacy represents an important aspect of teachers' professional knowledge.

## Research on assessment literacy

Language teachers with a solid background in assessment are well positioned to integrate assessment with instruction so that they utilise appropriate forms of teaching. Despite the importance that is given to being assessment literate, our progress toward an assessment-literate educational culture has been slow.

Research continues to characterise we teachers' assessment and evaluation practices as largely incongruent with recommended best practice (Galluzzo 2005, Mertler 2003, Zhang and Burry-Stock 1997 as cited in Volante and Fazio 2007:750).

In North America, there continues to be relatively little emphasis on assessment in the professional development of teachers. For example, out of ten Canadian provinces and 50 US states, only Hawaii and Nebraska currently invest significant funds which are specifically targeted to improve assessment and evaluation practices within schools (Volante and Fazio 2007).

Research on teaching in mainstream classrooms has revealed that the day-to-day assessment of student learning is unquestionably one of the teachers' most demanding, complex and important tasks (Calderhead 1996 as cited in Cheng 2001:54, Shulman 1986). Teachers view student evaluation as a central teaching function in their classrooms. This is evidenced by the time spent on assessment-related activities.

In the ESL education literature within North America, Bachman (2000) reported that a survey of the TESOL organisation membership conducted in the 1990s found about half of the respondents had completed a course in language testing and Stoynoff (2009) determined that about half of the graduate programs in the *Directory of Teacher Preparation Programs in TESOL* (Christopher 2005) required graduates to complete coursework in language assessment. These results are similar to a recent study completed by Brown and Bailey (2008) in which 60% of the respondents were from outside the US. Based on these data it appears half of all ESOL teachers may not have completed coursework in language assessment (Stoynoff and Coombe forthcoming).

While there is rich literature and a plethora of research studies on ESL/EFL teachers' assessment practices (e.g. Cheng, Rogers and Wang 2008) there continues to be a gap in the area of assessment literacy and what constitutes teachers' knowledge. In fact, as far as teacher preparation in assessment is concerned in EFL contexts, teachers in Hong Kong report that they received little or no training in assessment (Falvey and Cheng 1995). Shohamy (1998) and Ferman (1998) found that EFL teachers in Israel felt they lacked the knowledge and training required to practise assessment procedures. More recently, in a study done with tertiary-level English-language teachers in the United Arab Emirates and Kuwait, Troudi, Coombe and Al-Hamly (2009) found that teachers often felt marginalised in the area of assessment because of their perceived lack of knowledge about the subject.

## Defining Assessment Literacy

Interestingly, the term 'assessment literacy' is not listed in the *Dictionary of Language Testing* (1999), ALTE's *Multilingual Glossary of Language Testing Terms* (1998) or Mousavi's *Encyclopedic Dictionary of Language Testing* (2002). While each of these volumes devotes ample space to the concept of assessment, the issue of how educators become assessment literate is not mentioned. Despite the lack of definitions in these important assessment volumes, the term 'assessment literacy' has been defined by a number of well-known assessment experts.

According to Popham (2004) and Stiggins (2002) assessment literacy is simply an understanding of the principles of sound assessment. Implicit in this definition is that assessment literate teachers have the know-how and understanding needed to assess their students effectively and maximise learning.

Those educators who are deemed to be assessment literate are familiar with the principles of sound assessment and how to meet specific standards of quality. The characteristics of sound assessment according to Stiggins (2007) are that they:

1. arise from and serve clear purposes
2. arise from and reflect clear and appropriate achievement targets
3. rely on a proper assessment method (given the purpose and the target)
4. sample student achievement appropriately
5. control for all relevant sources of bias and distortion.

Assessment literate educators come to any assessment knowing what they are assessing, why they are doing so, how best to assess the achievement of interest, how to generate sound samples of performance, what can go wrong, and how to prevent these problems before they occur (Stiggins 1995:240). Language teachers and administrators need the necessary tools for analysing and reflecting upon test and assessment data in order to make informed decisions about instructional practice and program design.

By developing assessment literacy, language educators will not only be able to identify appropriate assessments for specific purposes, such as student placement, but will also be able to analyse empirical data to improve their instruction.

## Barriers to assessment literacy

There are a number of impediments or what Stiggins (1995) calls 'barriers' to assessment literacy.

The first and perhaps most important reason is 'fear'. According to Stiggins (1995), educators often carry with them an accumulation of layers of negative emotions associated with assessment. This fear of assessment has often been cultivated over many years of unpleasant assessment experiences. The foundations of this fear are often rooted in the assessments that we have undergone as young people. Fear represents a prominent barrier to assessment literacy because it closes many educators

off from even reviewing their own assessment competence.

Another reason why teachers do not want to become involved in or increase their knowledge in assessment is put forth by Alderson (2001). He states that the field of assessment is often viewed by teachers as an arcane 'Ivory Tower' where many of the journals are not accessible to the average classroom teacher.

Concerns close to the teachers' daily lives constitute another important reason for the lag in the development of assessment literacy. With the increasing demands of the workplace, some teachers feel that it is simply easier not to worry about assessment. These teachers are content to let others write the assessments for them.

Another significant barrier to assessment literacy is that there are insufficient resources allocated to assessment. It has been stated time and time again that although administrators pay lip service to the importance of assessment, very few actually back it up with the resources needed to make assessment programs more successful. Administrators view assessment and assessment-related activities as being part of a teacher's job and often do not provide reduced teaching loads or extra remuneration for those who get actively involved in such activities.

All of the factors mentioned above conspire against teacher involvement in assessment and increased levels of assessment literacy in our teachers.

## What assessment skills are needed to be assessment literate

A number of well-known assessment scholars and organisations have put forth lists of characteristics of what it takes to be assessment literate.

According to the *Seven Standards for Teacher Development in Assessment* developed by the American Federation of Teachers, the National Council on Measurement in Education and the National Education Association (1990), teachers should be skilled in:

1. choosing assessment methods appropriate for instructional decisions
2. developing appropriate assessment methods
3. administering, scoring and interpreting the results of both externally-produced and teacher-produced assessment methods
4. using assessment results when making decisions about individual students, planning teaching, developing curriculum and involving students
5. developing valid grading procedures which use student assessment
6. communicating assessment results to students, parents, and other stakeholders
7. recognising unethical, illegal and inappropriate assessment methods and uses of assessment information.

In a useful online publication from SERVE at the University of North Carolina, they recommend that assessment-literate teachers know:

- how to define clear learning goals, which are the basis of developing or choosing ways to assess student learning

- how to make use of a variety of assessment methods to gather evidence of student learning

- how to analyse achievement data (both qualitative and quantitative) and make good inferences from the data gathered

- how to provide appropriate feedback to students

- how to make appropriate instructional modifications to help students improve

- how to involve students in assessment process (e.g., self- and peer assessment) and effectively communicate results

- how to engineer an effective classroom assessment environment that boosts student motivation to learn. (SERVE Center, University of North Carolina, 2004)

Sadler (1998) shares these characteristics of an assessment literate educator:

- superior knowledge about content and substance of what is to be learned

- knowledge about learners and learning and a desire to help students develop, improve and do better

- skills in selecting and creating assessment tasks

- knowledge of criteria and standards appropriate to assessment tasks

- evaluative skills and expertise in the analysis and use of assessment information

- expertise in giving appropriate and targeted feedback.

In the TESOL/NCATE standards for ESOL teacher education, in the assessment domain, teachers are expected to understand issues of assessment for ESL and language proficiency assessment for ESL (including how to develop assessments and use them to inform instruction).

In short, those who are assessment literate understand what assessment methods to use in order to gather dependable information about student achievement, communicate assessment results effectively, and understand how to use assessment to maximise student motivation and learning.

## Recommendations for achieving assessment literacy

First, it is crucial that we develop a universal understanding of what constitutes a good assessment and to build a common, articulated set of criteria for exemplary language assessment. This certainly does not negate the recognition of different views about the nature of education which might lead to dissimilar approaches to assessment. There remains an urgent need to encourage and organise professional development through both online training of teachers and through assessment workshops at all levels.

If we are to achieve assessment literacy, we need to provide teachers with the requisite professional development and time to implement those practices learned. A few workshops are insufficient. Successful professional development in assessment will require significant change in our educational practices and a time commitment on the part of teachers.

Successful professional development in the area of assessment literacy needs to take into account the learning styles and workload of today's language teachers. In order for teachers to achieve assessment literacy the availability of assessment resources, especially online, is critical.

## Conclusion

Teachers will be expected to be far more assessment literate in the future than they are today or have been in the past (Stiggins 2007).

Assessment literate educators come to any assessment knowing what they are assessing, why they are doing so, how best to assess the achievement of interest, how to generate sound samples of performance, what can go wrong, and how to prevent these problems before they occur (Stiggins 1995:240).

It is best stated by Bracey (2000), there might come a time when tests and test scores recede from prominence, but that time is not now. In view of the importance of assessment in today's educational institutions around the world, 'assessment literacy' is a necessity for all language educators.

## References and further reading

Alderson, J C (2001) Testing is too important to be left to testers, in Coombe, C (Ed.) *Selected Papers from the 1999 and 2000 CTELT Conferences*, Dubai: TESOL Arabia Publications, 1–14.

ALTE (1998) *Multilingual glossary of language testing terms*, Studies in Language Testing volume 6, Cambridge: Cambridge ESOL/Cambridge University Press.

American Federation of Teachers, National Council on Measurement in Education, National Education Association (1990) Standards for teacher competence in educational assessment of students, *Educational Measurement: Issues and Practice* 9/4: 30–2.

Bachman, L F (2000) Modern language testing at the turn of the century: Assuming that what we count counts, *Language Testing* 1/1, 1–42.

Bracey, G (2000) *Thinking about tests and testing: A short primer in assessment literacy*, paper presented at the American Youth Policy Forum, Washington, DC, ED 445 096.

Brown, J D and Bailey, K M (2008) Language testing courses: What are they in 2007? *Language Testing* 25/3, 349–83.

Calderhead, J (1996) Teachers: Beliefs and knowledge, in Berliner, D C and Calfee, R C (Eds) *Handbook of educational psychology*, New York, NY: MacMillan Library Reference, 709–25.

Cheng, L (2001) An investigation of ESL/EFL teachers' classroom assessment practices, *Language Testing Update* 29, 53–83.

Cheng, L, Rogers, T and Wang, X (2008) Assessment purposes and procedures in ESL/EFL classroom, *Assessment and Evaluation in Higher Education* 33: 9–32.

Christopher, V (2005) *Directory of teacher education programs in TESOL in the United States and Canada*, Alexandria, VA.: TESOL.

Cowan, J (1998) *On becoming an innovative university teacher*, Buckingham: RHE and Open University Press.

Davies, A, Brown, A, Elder, C, Hill, K, Lumley, T, and McNamara, T (1999) *Dictionary of language testing*, Cambridge: Cambridge University Press.

Falvey, P and Cheng, L (1995) A comparative study of teachers'

beliefs about assessment principles and practices, *Language Testing Update* 18, 38–9.

Ferman, I (1998) *The impact of a new English foreign language oral matriculation test on the educational system*, unpublished MA thesis, Tel Aviv University.

Galluzzo, G R (2005) Performance assessment and renewing teacher education, *Clearing House* 78/4, 142–45.

Herman, J and Dorr-Bremme, D (1982) *Assessing students: Teachers' routine practices and reasoning*, paper presented at the annual meeting of the American Educational Research Association, New York, NY.

Jacobs, L C and Chase, C I (1992) *Developing and using tests effectively: A guide for faculty*, San Francisco, CA: Jossey-Bass.

James, R, McInnis, C, and Devlin, M (2002) *Assessing learning at Australian Universities*, Center for the Study of Higher Education, the University of Melbourne, Australia, available online http://www.cshe.unimelb.edu.au/assessinglearning/

McLaughlin, P and Simpson, N (2004) Peer assessment in first year university: How the students feel, *Studies in Educational Evaluation* 30/2, 135–49.

Mertler, C (2003) *Preservice versus inservice teachers' assessment literacy: Does classroom experience make a difference?* Paper presented at the annual meeting of the Mid-Western Educational Research Association, Columbus, OH, October.

Mousavi, S A (2002) *Encyclopedic dictionary of language testing* (3rd Ed.), Taipei: Tung Hua Book Company.

Popham, W J (2004) All about accountability: Why assessment illiteracy is professional suicide, *Educational Leadership* 62/1, 82–3.

Sadler, D R (1998) Formative assessment: Revisiting the territory, *Assessment in Education* 5, 77–84.

SERVE Center (2004) Classroom assessment: Assessment literacy, University of North Carolina, available online http://www.serve.org/Assessment/Classroom/Literacy.php

Shohamy, E (1998) Inside the 'black box' of classroom language tests, *Studia Anglica Posnaniensia* XXXIII, 343–52.

Shulman, L S (1986) Paradigms and research programs in the study of teaching: A contemporary perspective, in Wittrock, M C (Ed.) *Handbook of research on teaching* (3rd Ed.), New York, NY: MacMillan.

Stiggins, R J (1995) Assessment Literacy for the 21st Century, *Phi Delta Kappan* 77/3.

– (2002) Assessment crisis: The absence of assessment for learning. *Phi Delta Kappan* 83/10, 758–65.

– (2007) Conquering the formative assessment frontier, in McMillan, J (Ed.) *Formative Classroom Assessment*, New York, NY: Colombia University Teachers College Press, 8–28.

Stiggins, R J and Conklin, N (1992) *In teachers' hands: Investigating the practice of classroom assessment*, Albany, NY: SUNY Press.

Stoynoff, S (2009) A survey of developments in ESOL testing, in Coombe, C, Davidson, P and Lloyd, D (Eds) *Fundamentals of language assessment: A practical guide for teachers*, Dubai: TESOL Arabia Publications.

Stoynoff, S and Coombe, C (forthcoming) Professional Development in Language Assessment, unpublished manuscript.

TESOL/NCATE Program Standards (2003) available online http://clas.uncc.edu/linguistics/Internal%20documents/NCATEP12Standards.pdf Alexandra, VA: TESOL.

Troudi, S, Coombe, C and Al-Hamly, M (2009) EFL teachers' views of English language assessment in higher education in the United Arab Emirates and Kuwait, *TESOL Quarterly* 43/3, 546–55.

Volante, L and Fazio, X (2007) Exploring teacher candidates' assessment literacy: Implications for teacher education reform and professional development, *Canadian Journal of Education* 30/3, 749–70.

Wise, A E (Ed.) (1996) Quality teaching for the 21st Century (special issue), *Phi Delta Kappan* 78, 190–224.

Zhang, Z and Burry-Stock, J (1997) *Assessment practices inventory: A multivariate analysis of teachers' perceived assessment competence*, paper presented at the Annual Meeting of the National Council on Measurement in Education, Chicago, IL.

# Dimensions of teacher development in a Romanian higher education context

**LAURA MUREŞAN** BUCHAREST UNIVERSITY OF ECONOMICS, ROMANIA

## Introduction

Increased mobility for academic and work purposes, the trend towards internationalisation of higher education programmes and a stronger focus on quality standards in education, in general, bring about new Teacher Development needs and interests and, at the same time, open up new opportunities for language professionals. With the Romanian universities' commitment to develop English-, French- or German-medium masters programmes, for example, there are new roles evolving for language teachers and teacher educators in the field of languages for special professional and academic purposes (LSP/ESP-EAP). There is a lot of scope for their diversifying and

enriching existing competencies, as well as for transferring their expertise and methodological approaches to other domains. There is also the question of finding the best possible ways for coping with change and turning challenges into opportunities.

This article comprises two Teacher Development case studies from the Romanian Higher Education context. The first one will illustrate how individual interests and needs can be integrated in an institutional context through an eclectic approach to teacher development and mentoring at the Department of Business English and German of the Bucharest University of Economics. The second is an example of transferring expertise developed through ESP-EAP teacher education to other domains within the

same university, in the structured framework provided by an English-medium interdisciplinary Research and Teacher Education Masters programme, designed and co-ordinated by ESP-EAP teacher trainers for academics and researchers of various specialisations.

## The Professional Development and Mentoring Programme at the Department for Business English and German, the Bucharest University of Economics

This programme was designed and introduced by an academic community of language professionals to the benefit of the individuals participating in the programme, as well as that of the organisation, in this case, the Department for Business English and German of the Bucharest University of Economics. This example will allow us to explore various dimensions of personal and institutional learning, such as motivation, self-evaluation, inter-generational sharing, the importance of a positive, collegiate atmosphere, and the inter-dependencies between professional development and quality enhancement.

The programme was initiated at the end of 2004, in a particular stage in the Department's development, when the number of staff had almost doubled as compared to previous years. As in-house teacher training events introduced as induction for new staff members at the beginning of the academic year had received positive feedback, it was felt that a smooth transfer of expertise was not only needed but would also be welcomed by all of the teachers – both junior teachers and more experienced peers, as discussed below.

### Main programme aims

From the outset, the programme aims were intended to address various levels, from the individual to the institutional, both short-term and medium-term. The main aims included:

- facilitating the professional development of junior staff members
- providing a framework for multiplying the opportunities for sharing experience and working together
- integrating the professional expertise of each member within a coherent 'continuum' that would contribute to the personalised development of all those involved, as well as to team development
- familiarisation with best practice models for teaching and (self-)evaluation, thus contributing to better methodological coherence in view of improved quality management at department level
- capacity building, consolidating the institutional profile of the Business English and German Department as organiser of post-graduate studies, including teacher training for specific fields – e.g. at Master's level (medium-term), at PhD level (long-term).

In addition, medium-term, if the programme proved its usefulness, wider objectives also included the transfer of expertise to other departments of the same university or even to other educational contexts or institutions.

After a brief outline of the programme participants' profile, the main stages of this Teacher Development process are presented in the following setions.

### The participants in the programme

*Teachers of business English and German*

The main beneficiaries of the programme were less experienced teachers of business English and business German, junior assistants who had joined the Department one or two years prior to the programme's start. All of them were Philology graduates, with a high level of proficiency in the target language, most of them with some experience of teaching general English or German but with fairly limited experience of teaching specialised language courses for business or various economic specialisations. Most importantly, all of them had great enthusiasm and thedesire to develop specific teaching skills, to improve their knowledge of specialised English (German) for economics, and to become more confident in assessing student performance.

*Mentors*

The teachers taking on a mentoring role were experienced ESP-practitioners, with teacher training and mentoring expertise, developed through the British Council PROSPER-Project (Bardi 1999), as well as through quality management courses within the framework of QUEST Romania (The Romanian Association for Quality Language Services) and EAQUALS (The European Association for Quality Language Services). In addition, they were building on project management, materials writing and facilitation competencies developed through participation in European projects, e.g. Quality Assurance and Self-assessment for Schools and Teachers and QualiTraining within the framework of the ECML – the European Centre for Modern Languages of the Council of Europe. And last but not least, one of their key characteristics was that of being good communicators, happy to share and learn by doing, not afraid to learn from mistakes and to talk about them.

*Institutional partners*

The main institutional partners who contributed to this programme in various stages of its unfolding were the British Council Romania, the Goethe-Institut Bucharest, the PROSPER-ASE Language Centre, and the Romanian Association for Quality Language Services QUEST Romania. Guest speakers and trainers included academics and professionals from abroad, as exemplified below.

### Planning, launch and needs analysis

The planning and preparation stage included consultation meetings with the Dean, with a British Council project manager (former coordinator of the PROSPER Project), consultation with experienced teachers who had developed relevant mentoring competencies, as mentioned above (Bardi et al. 1999, Bardi 2007), as well as informal discussions with junior teachers, who had recently joined the department.

The programme as such was launched at the beginning of 2005 in a meeting involving both experienced and less experienced teachers of business English and German. The presentation of the programme included an outline of the framework and main objectives, examples of possible themes to be covered or taken up for study, an induction to the underlying principles of how such a process could work to the benefit of all those involved, an outline of the format, with examples of possible activity types; all senior staff members (i.e. experienced teachers and mentors) were invited to briefly present themselves with a focus on their key strengths and areas of interest, so as to facilitate the junior teachers' formulating preferences regarding mentoring relationships. The Dean helped further contextualise this type of endeavour against the background of developments in Higher Education at European level, highlighting the role continuing professional development could have also for doctoral studies and for a teaching career in Higher Education.

The less experienced teachers were invited to fill in a questionnaire aimed at inducing a reflective approach to their own career plans and priorities, as well as to their strengths and areas for improvement, their professional interests and further suggestions regarding this programme (see Figure 1). They were also invited to indicate three options for mentoring, based on both the mentors' areas of expertise and 'chemistry'. All the 20 junior teachers participating in the meeting responded, thus contributing to the needs analysis exercise and to the forming of small, informal 'mentoring groups', each consisting of one experienced teacher, in an advisory role, and 1–3 junior teachers.

**Figure 1: Questionnaire for the Participants in the Professional Development and Mentoring Programme**

**Questionnaire**

1. Where would I like to be in 5 years' time?
2. Language courses that I have already taught (including reference to the year of study, language level, etc.)
   a. Strengths I feel I have
   b. Areas for improvement that I've identified
3. Courses and seminars that I would like to teach
4. Areas of interest
   a. where I already have expertise
   b. where I would like to study more
5. How do I learn best?
6. Who would I like to work with as a Mentor?
   Option 1 _____  Option 2 _____  Option 3 _____
7. Further suggestions re: topics to be covered

Once the framework was created, the programme started developing in a semi-structured, eclectic form, as illustrated in the next section.

### Format of events and thematic areas

What characterised this programme was a combination of events in various formats, such as:

- teacher training and development sessions on themes of interest, based on the initial needs analysis

- class observation, carried out as peer observation, followed by feedback discussions, peer-review, often combined also with peer-teaching
- mentoring (either one-to-one or in small groups), depending on thematic interest and expertise, as well as on 'chemistry'
- round-table discussions with guest speakers
- project events (e.g. within the framework of the Lingua 1 *EuroIntegrELP* project)
- team work on materials writing (sharing and learning 'on-the-job')
- conference participation and presentations
- self-development according to personal professional interest, in several cases leading to the participants starting their doctoral study.

One of the main thematic areas was *Implementing self-assessment in the teaching and learning process*, based on *The Common European Framework of Reference for Languages* (Council of Europe 2001) and the *European Language Portfolio* (EAQUALS-ALTE 2000) – a complex process, within the framework of the Lingua 1 *EuroIntegrELP* project. This consisted of input sessions and studying relevant materials (e.g. Brown 2005, Heyworth 2004), piloting of instruments and activities in class, followed by the sharing of experience in workshop format. In addition, most of the participants contributed to a feasibility study coordinated by the Romanian Institute of Educational Sciences on introducing the ELP in language education, as well as to a survey on the effectiveness of ELP implementation. An important outcome consisted of developing materials to integrate ELP-based self-assessment in the business English/German curriculum and the business English/German communication class.

Other themes addressed included:

- The Relevance of Critical Thinking for Business Communication
- Continuing Professional Development in UK universities and the inter-relatedness with Quality Assurance (Luxon 2006b)
- Induction to the EAQUALS Self-assessment Grid for Language Teachers with descriptors for teacher competencies (North and Mateva 2005, North 2009)
- Critical Discourse Analysis (CDA) presentations for all the participants and, over a period of almost three years, participation of those interested in the CDA Research Group *Re-scaling Romania*, coordinated by Norman Fairclough, Professor Emeritus with Lancaster University (Fairclough 2006)
- Academic Research and Requirements of Scholarly Publication (Ives and Obenchain 2007)
- Integrating professional and communication skills – an advanced, intensive course for teachers of business English, in preparation for the LCCI exam leading to the LCCI Further Certificate for Teachers of Business English; this was followed by most of the participants taking the LCCI FTBE exam and receiving the LCCI Certificate for Teachers of Business English

- Quality Management in language education and aspects of educational management in language teaching institutions/departments-based on the ECML projects on Quality Management in language education (Mureşan et al. 2003).

As we can see, content wise, a whole range of themes were addressed, while integrating individual plans in the overall programme framework, moving from the general to specific events and activities, establishing priorities, at the same time identifying who could make valuable contributions on topics of interest and finding ways of inviting guest speakers.

The formal, more structured part of this professional development programme totalled over 60 contact hours (input sessions, workshops, round-table discussions), as exemplified above. In addition to these, there were numerous instances of class observations, pair- or small-group discussions, team work on joint projects, and last but not least, informal communication, networking, informal exchanges, simply enjoying the sharing of impressions and valuing the time spent together, irrespective of age or teaching experience.

Although initially intended for three semesters, as a result of the participants' interest in pursuing their professional goals and also thanks to opportunities of inviting and benefiting from relevant contributions by external experts, the programme was extended till the end of 2007, thus covering almost three years, with a positive impact also on further developments in various areas, as shown below.

### Programme outcomes and effects

A proper, formal evaluation of the entire programme and its outcomes has not been carried out yet, therefore, evaluative remarks are mainly based on informal feedback, as well as on effects such as the participants' wish to extend the programme time-frame. So far, there have been evaluation exercises carried out for programme components, such as a survey among programme participants, initiated by the programme coordinator, regarding the implementation of self-assessment. In addition, at the end of 2007, following an external evaluation of the *EuroIntegrELP* project by the National Agency for Educational EU Programmes in Romania, with a special focus on Teacher Training and Development (of which the ELP-based TD component of this programme was an integral part), the EuroIntegrELP project received the *European Label Award for Innovation in Language Education*.

Concrete outcomes of the TD and Mentoring programme include, for instance, chapters on ELP-based self-assessment and on EUROPASS, integrated in several of the Business English textbooks produced by members of the Department.

At an institutional level, the programme was included in the Department's Professional Development Programme and counted as a plus in the quality assurance audit.

Other effects, more difficult to pinpoint as direct programme outcomes, could be seen in areas such as: co-authoring of business English textbooks, where the teams unite programme participants irrespective of 'generation';

the enrolment and participation of several participants in the interdisciplinary Research and Teacher Education Masters programme intiated in September 2006; the friendly, co-operative atmosphere in the department, with people enjoying the sharing and exchanging of experience, impressions, concerns. 'Experience' is no longer an issue, all the teachers have relevant teaching and assessment expertise, most of them were promoted to lecturer position and started teaching also at Masters level. 'Generation' is no longer an issue either. Joint participation in various project events and conferences, outside the work-related environment, definitely contributed to a natural process of team-building, without the artificial involvement of external team-building facilitators. Luxon (2006a:8) states:

> 'It is quite clear that there is a great deal of commitment to professional development among the ELT community in Bucharest, and that there is a generational quality relating to those who were involved in the Prosper project, and the younger teachers who have been influenced by the approaches developed through the project. This seems to be a great strength which needs to be exploited.'

Future plans include the design and carrying out of an impact study among all the stakeholders, and at the same time, continuing to explore possibilities for formal institutional recognition of the participants' professional development. Areas of interest signalled by the participants include research methodology, refining writing competencies for scholarly publication, project design and management, thus taking further, at another qualitative level, professional development and quality enhancement.

Some of these plans are intertwined with an important offspring of this TD programme: the interdisciplinary Masters, English Language Education and Research Communication for Business and Economics, briefly presented in the next section.

## Interdisciplinary Research and Teacher Education Masters Programme at the Bucharest University of Economics

The university's interest in introducing English-medium modules in bi-lingual programmes and new programmes, entirely in English, thus encouraging subject teachers to switch from teaching their specialisation in their L1 to doing this in English, has brought about new challenges. At the same time, the university's opting for the profile and status of a research university has put additional pressure on academics, who are urged to publish in international research journals.

Thus, three years ago (in the fall of 2006), a new interdisciplinary Masters programme was introduced at the Bucharest University of Economics, in response to both institutional and individual needs. 'English Language Education and Research Communication for Business and Economics' has elements both of a Research MA and an M.Ed., aiming to address needs in both areas.

Worthwhile mentioning is that this masters programme was designed and is co-ordinated by the same ESP teacher trainers who initiated the TD and Mentoring programme presented above, transferring expertise from ESP to other subject fields in Higher Education. In addition, it is an

illustration of how ESP-TD can be further diversified and refined to address new areas, beyond language improvement, taking on new dimensions, conducive to the professional development of experts in other fields.

### Main programme aims

This interdisciplinary Research and Teacher Education Masters programme is specifically intended to contribute to:

a. improving teaching methodology and communication competencies, with a focus on CLIL in Higher Education for business and economics

b. facilitating the participants improving their English language proficiency, for both teaching and research purposes

c. enhancing the quality of research in terms of methodological aspects, while integrating professional and English language communication skills for the successful presentation of research outcomes in international events and specialised publications.

### Target participants

This programme brings together professionals with an interest in improving their teaching and research methodology, irrespective of their professional background. Thus, the main categories of candidates include:

• academics of various specialisations (e.g. business management, micro-economics, marketing, agri-business, finance, accounting, cybernetics, but also engineering, mathematics, economic geography, history, etc.) already teaching or interested to teach their subject through the medium of English

• researchers interested in the interdisciplinary dimension and in improving their English language and communication competencies for publishing research outcomes internationally and for participation in international projects

• philology graduates teaching English for business and economics (or with an interest in ESP teaching and research)

• graduates of various specialisations with an interest in the programme objectives and components.

With such a range of backgrounds and interests, the transversal dimension of the programme is achieved both through the content-components of the curriculum and the methodological approach promoted, as illustrated in what follows.

### Thematic areas and methodological approach

The interdisciplinary dimension, the curriculum and the methodological approach, are meant to let the participants experience a different way of teaching, learning and (self-)evaluation. Networking with colleagues of other specialisations is encouraged, as well as thinking outside the box, taking the challenge of exploring aspects outside one's own subject field. One of the outcomes – an added value of this programme – is that of achieving more consistency and coherence of approach in terms of quality assurance across disciplines.

Content-wise, the programme includes (a) modules with a teaching focus, (b) modules with a research focus, (c) modules aimed at developing advanced communication skills for a variety of academic and professional purposes, as well as transversal components, e.g. on 'Critical Thinking' and 'Creative Thinking', on 'Educational Management' and 'Quality Assurance in Higher Education', etc.

An important element of this interdisciplinary Masters is the reflective and self-evaluative component. Thus, key to progress and the programme's effectiveness is the methodological approach, which includes e.g.:

• reflection on one's own teaching context and activity, with a focus on skills improvement and prioritising self-development objectives

• self-assessment of communication skills, on the basis of the European Language Portfolio and more detailed, relevant reference scales of levels from the Common European Framework of Reference, both in relation to one's own targets and in relation to programme requirements (as this is an English-medium programme)

• reflection on institutional processes through team-exercises, aimed at improving self-evaluative competencies

• interdisciplinary project work and assignments encouraging reflective practice

• observation of processes, including class observation and peer-review.

The impact of this methodological approach can be seen at various levels, for instance in the selection of research projects for dissertations. For example, a good number of participants select topics involving qualitative analysis of student perceptions of teaching or assessment related methodological aspects, quality assurance aspects from a practical, grassroots perspective (a synthetic version of several dissertations was published in Synergy, volume 4/2, 2008, e.g. Ion 2008, Serban-Oprescu 2008, Zaharia 2008).

In their feedback, most of the participants in the programme indicate that their expectations had been exceeded, especially through the interdisciplinarity, the networking opportunities, the international dimension of the programme, the reflective and self-evaluative competencies that they developed, and the impact these had on their teaching and research practice. This way, self-evaluative approaches contribute to improved performance of individual teachers, as well as to sharing experience and team learning.

As multipliers and having had the chance to experience these methodological approaches themselves, most of the participants enjoy taking these further into class, to their students. This way, in a cascading mode, we can also witness a shift of paradigm, with teachers developing the confidence 'to let go' and to encourage students to voice their preferences, engaging them in new processes.

## Conclusion

If we take the perspective of ESP teachers participating as students in this Masters programme, it means opening up new interdisciplinary perspectives and an increased

awareness of what is needed in the 'real world', including the 'real academic world' typical of other subject fields in business and economics. For all the participants in the programme – irrespective of their subject field – it also means skills development in areas such as critical thinking and creative thinking, long-term thinking and forecasting developments in their field by taking into account emerging trends in general, as well as advanced EAP skills, which they can then apply in their own ESP classes; practical aspects of educational management and quality assurance (e.g. the module based on the ECML *QualiTraining* project – Mureşan, Heyworth, Mateva and Rose 2007), applicable to their own teaching context; experiencing interdisciplinary project work, usually taken further also outside the programme framework and multiplied in work with their own students, thus adding new dimensions to ESP learning.

If we take the perspective of the academics involved on the teaching/training side, it is definitely a challenging experience, at the same time rewarding and confidence-building. As we can see also from other academic contexts, ESP/EAP teacher trainers can bring their contribution to the transfer of expertise relevant also to other specialisations and, thus, to the quality enhancement of educational processes within the university.

The two case studies from Romanian Higher Education presented above illustrate various stages on the teacher/trainer development continuum: from an interest in improving one's own teaching and assessment methodology, to developing new skills (e.g. for authoring textbooks, conference presentations, etc.), moving on to consolidating teacher development and mentoring competencies through learning partnerships, and further diversifying and refining TD competencies, transferring relevant aspects developed for and through ESP-EAP teacher education to other domains.

## References and further reading

Bardi, M, Chefneux, G, Comanetch, D, Magureanu, T (Eds) (1999) *The PROSPER Project. Innovation in Teaching English for Specific Purposes in Romania – A Study of Impact*, Bucharest: The British Council Cavallioti.

Bardi, M (2007) *Managing Innovation – Aspects of Professional Learning*, Bucharest: Cavallioti.

Brown, P (2005) *The Future – the eELP (electronic European Language Portfolio)*, paper presented at the Joint QUEST-RATE Conference: Evaluation – A Requisite for Quality Learning, Cluj-Napoca, 28–30 October 2005.

Council of Europe (2001) *Common European Framework of Reference for Languages: Learning, teaching, assessment*, Cambridge: Cambridge University Press.

EAQUALS-ALTE (2000) The *EAQUALS-ALTE European Language Portfolio*, Milano: Lang Edizione available online www.eELP.org

Fairclough, N (2006) *Language and Globalisation*, London and New York: Routledge.

Heyworth, F (2004) Why the CEF is Important, in Morrow, K (Ed.) *Insights from the Common European Framework*, Oxford: Oxford University Press, 12–21.

Ion, R A (2008) A New Approach to Teaching and Learning, *Synergy* 4/2, 48–66.

Ives, B and Obenchain, K (2007) Scholarly Publication, *Synergy* 3/1, 7–14.

Luxon, T (2006a) *Report on a Consultancy Visit to Bucharest on behalf of the British Council by Dr. Tony Luxon of Lancaster University*, unpublished document.

– (2006b) Reflections on the Reflective Practitioner: Investigating our own practices, *Synergy* 2/1, 92–103.

Mureşan, L, Heyworth, F, Matheidesz, M and Rose, M (2003) *Quality Management in Language Education*, Strasbourg: Council of Europe Publishing, CD ROM, available online http://www.ecml.at/mtp2/qualitraining/results/quality/index.htm

Mureşan, L, Heyworth, F, Mateva, G and Rose, M (2007) *QualiTraining – A Training Guide for Quality Assurance*, Strasbourg: Council of Europe Publishing, available online www.ecml.at

North, B (2009) *A Profiling Grid for Language Teachers*, paper presented at the International Meeting on Training, Quality and Certification in Foreign Language Teaching, Siena, 9–10 February 2009, available online www.eaquals.org

North, B, Mateva, G (2005) *A Professional Profile for Teachers: Is a Teacher Portfolio Feasible?* paper presented at the EAQUALS Conference: European Quality Standards – Accessing and Assessing Quality, Athens, 24–26 November 2005.

Serban-Oprescu, A T (2008) Project Management and Effective Professional Development in Higher Education, *Synergy* 4/2, 172–91

Zaharia, R M (2008) Performance of Academic Research. Case Study: Bucharest University of Economics. The View of Academics, *Synergy* 4/2, 26–30.

# Addressing teacher needs: how Cambridge ESOL's teacher support and professional development are being expanded

**ANDREW NYE** STAKEHOLDER RELATIONS GROUP, CAMBRIDGE ESOL
**KAREN BARNS** STAKEHOLDER RELATIONS GROUP, CAMBRIDGE ESOL

## Introduction

Cambridge ESOL's global teacher community is large and varied; in addition to obvious cultural differences, the level of affluence, resources and IT infrastructure differ considerably across the world. In order to support this diverse community, Cambridge ESOL currently offers a range of products which fall into three categories: printed; face-to-face; and online support. In March 2009 an online questionnaire was deployed to gather opinions from across this community as to how Cambridge ESOL's training and support could be developed further. This article looks at what Cambridge ESOL currently offers the teaching community and, with reference to the feedback from the teacher questionnaire, what possibilities there are for future expansion.

## Background to the Teacher Survey

The online questionnaire was sent to just over 10,000 teachers, of which 2119 responded (a response rate of 21%). The respondents were largely European, 47% coming from Western Europe. The next largest group were from South America (27%) and the remainder from Eastern Europe (6%), the UK (6%) and Asia (2%).

The majority of the sample was experienced teachers: 75% of them had been teaching for more than 5 years, and 60% for more than 11 years. However, only 25% of the sample was native speakers, the majority of 57% rating themselves as having 'advanced' level English language ability. Although experienced, 72% of the teachers said they would like a preparation course for TKT, which suggests the sample as being largely untrained.

The teachers worked in a variety of sectors: language schools (35%), secondary schools (32%), universities (16%), and primary schools (8%). Half of the respondents taught classes of less than 15 students and 28% had between 16–24 students in their classes. A minority of 7% taught classes of 35 or more.

All of the teachers were currently preparing students for a Cambridge ESOL exam and many were preparing for more than one level of Cambridge ESOL exam. The majority (64%) taught FCE, 39% prepared students for CAE, (also 39% for PET), 30% taught KET classes, and 25% Proficiency. 17% of the sample taught Young Learners, which was the same figure as IELTS. Of those teaching in the UK, 64% taught Skills for Life.

## Printed support

Cambridge ESOL's current portfolio of printed teacher support consists of:

- newsletters, e.g. *Cambridge First* (provides up-to-date news, views and developments in Cambridge ESOL's exams and Teaching Awards and how they are used around the world) and *Research Notes*

- exam handbooks: for anyone preparing candidates for Cambridge ESOL exams, each handbook provides:
  - an overview of the exam and its place within Cambridge ESOL
  - the full specification for the exam
  - guides to each of the papers or components
  - full sample papers
  - detailed information on how the exam is assessed.

- exam reports for KET, PET, FCE, CAE, CPE and YLE give:
  - an overview of candidate performance in each of the exam papers
  - tips for effective candidate preparation.

- past papers: full sets of question papers for different exam sessions are available in a booklet with mark schemes and answer keys. They are accompanied by an audio CD of the Listening test.
  N.B. Past papers are now being made available in an alternative format of classroom sets of ten past papers including mark scheme, answer key and listening CD, together with speaking visuals

- Speaking Test Preparation Packs
  - designed to help teachers prepare students for the Speaking test, with clear explanations of what each part of the test involves and step-by-step guidance and practical exercises
  - contain comprehensive Teacher's Notes, photocopiable Student Worksheets and (where relevant) colour versions of the candidate visuals
  - with an accompanying DVD containing video footage of students taking the tests.

As can be seen, the current range of printed support is broad. However, response from the survey suggests there is a demand for Cambridge ESOL to produce more support for classroom preparation. Speaking Preparation Packs are already produced for KET for Schools, PET for Schools, FCE, CAE, CPE, BEC Preliminary, BEC Vantage and BEC Higher. This support is clearly highly valued, as 95% of the teachers said they would be interested/very interested in also having

preparation packs for Writing, Use of English, Reading and Listening papers.

Work has started on designing the concept of Writing Test Preparation Packs. These will contain samples of candidate writing, (in place of the video clips used in the Speaking Packs) and will be based on the following core ideas:

- assessing the level of a piece of student written work. What is it that makes the writing the level that it is? What are its strengths and weaknesses? Key to understanding how writing is assessed will involve looking at the importance of range, accuracy, organisation and cohesion

- demonstrating the standard a piece of writing needs to get to, to be of the required level for the exam in question

- identifying what the priorities are, and how to address those priorities, in order to get the writing to that required standard.

Future projects will look at producing Preparation Packs for the other skills.

## Face-to-face training courses

Currently Cambridge ESOL offers an extremely popular seminar programme, covering all of Cambridge ESOL's exams and a wide range of topics (e.g. Recognition, European Language Portfolio, Speaking Assessment Scales). Trainers for these seminars are selected because of their training skills and experience of teaching Cambridge ESOL exams. The seminars are written by Cambridge ESOL and are therefore based on the latest thinking behind the exams and provide the most up to date information. However, trainers are able to adapt the seminars to their local audience, ensuring that the sessions are always relevant.

In response to the success of these seminars, teachers are asking for more training covering a wider range of topics. The Teacher Survey highlighted a demand for teacher training which goes beyond Cambridge ESOL exams. Teachers in the survey were asked if they would be interested in the teacher training workshops in the subject areas listed in Table 1. The percentage shows how many of them would be interested/very interested in taking these courses:

**Table 1: Percentage of teachers who expressed an interest in taking a course on various topics**

| | |
|---|---|
| correcting spoken and written work | 85% |
| assessing learners | 78% |
| strategies of supporting learners in mixed level classes | 76% |
| classroom management | 65% |
| planning a sequence of lessons | 61% |
| classroom language | 61% |
| planning and structuring a lesson | 60% |
| using the course book | 55% |
| use of local language in the classroom | 52% |
| strategies of supporting learners in classes of 30 or more students | 47% |

Cambridge ESOL has already begun extending the level of face-to-face support it delivers. One example is the course earlier this year for the Thai Ministry of Education, in partnership with Bell International. A group of 30 English

Language teachers from Thailand were given the opportunity to enhance their skills on a four-week training course in Cambridge.

The Thai Ministry of Education had selected teachers who had scored highly in the Teaching Knowledge Test to take part in the course. Working closely with the Thai Ministry of Education, Cambridge ESOL and Bell International delivered the training to help delegates develop their knowledge of teacher training methodologies and improve their overall skills as practicing teachers. The course also equipped the teachers with the necessary skills to design and develop their own English language tests.

Initiatives such as the course for Thai teachers are set to continue, but Cambridge ESOL is currently investigating how it could bring training to the wider teaching community. As Mike Milanovic, CEO for Cambridge ESOL commented: 'Providing teachers from all over the world, with all of the necessary support and guidance when educating and assessing students who are learning a foreign language is a key objective for us'.[6]

The challenge for Cambridge ESOL is how to bring its valued support to a worldwide audience. Face-to-face support is undoubtedly well-received and valuable, but due to both time and geographic constraints some teachers are unable to attend seminars. In the Teacher Survey, 90% of the teachers interested in training said they would like to receive it either as an online course, or as a combination of face-to-face and online training. It is this area of Cambridge ESOL's support which is looked at next.

## Online support

Cambridge ESOL currently offers the following online support:

- The Cambridge ESOL website: www.cambridgeesol.org/index.html This provides general information about Cambridge ESOL and the exams it offers.

- The Teacher Portfolio website: www.teacherportfolio.cambridgeesol.org/ This is a free online system that any teacher can use to record and document their career progress. It is designed to be a tool for life long learning and professional development.

- The Teacher Support website: www.teachers.cambridgeesol.org This recently launched site has replaced the previous Teaching Resources Website (www.cambridgeesol.org/teach/).

Feedback from the Teacher Survey suggests that the new Teacher Support website is a very welcome addition to Cambridge ESOL's teacher support package. Although the Teaching Resources website provided valued classroom support, teachers commented that it was very static and lacked the provision for teachers to be able to interact with both Cambridge ESOL and other teachers around the world.

The new Teacher Support site addresses this issue by giving teachers the opportunity to contribute and become directly involved with the teaching and examining

---

6   News item dated 28 July 2009 online at www.cambridgeesol.org

community. It also enables the organisation to listen to what teachers need, by providing forums for feedback and discussion for every Cambridge ESOL exam. This new Teacher Support site contains:

- A home page with a search facility and a registration feature.

- An overview of Cambridge ESOL exams with information about the papers, materials to download and testimonials. This includes all the exams under the categories of General English, Business English, English for Schools (including YLE) and Academic English. There is also an area for Teaching Qualifications, (TKT, CELTA, DELTA etc.).

- A resources area to house materials created by both Cambridge ESOL and teachers. Here teachers can download lesson plans written by Cambridge ESOL which can be used to introduce the exams to their students. In addition to an overview activity, there are activities to introduce each part of a paper. There is a search facility which enables teachers to find activities quickly and easily, and teachers are able to rate and comment on the resources they download.

- An area called Advice for Teachers which offers teaching tips and strategies for every paper of each exam. There are also FAQs at this level.

- An events area where teachers can look for events happening in their country by exam.

- Moderated discussion forums. As already mentioned, this key feature of the new site enables teachers to interact with other teachers and Cambridge ESOL on a new level.

- Links to exams dates.

- Links to publishers and published materials by exam.

The new Teacher Support website is the first in a series of online initiatives being planned by Cambridge ESOL. Cambridge ESOL is aiming to expand its teacher support by offering continuing professional development as part of a teacher's everyday life, removing the need to travel to an institution or be confined to a specific schedule. To achieve this, discussion is taking place as to how more online

training can be provided. Short online courses to complement the existing face-to-face seminar programme is one possibility being considered. Such courses would enable teachers to work towards their developmental goals by taking a pick and mix assortment of courses relevant to their individual interests and needs.

## Other initiatives

The results of the Teacher Survey have shown the value of Cambridge ESOL getting to know its teaching community better. To this end, work is underway to develop the concept of teacher relationship management, and more specifically a central teacher database which will enable Cambridge ESOL to consistently distribute messages and materials more widely and more directly to its teachers. It will enable the organisation to segment the huge teacher community into profiles (e.g. geographical area, organisational type, age, experience) and target those profiles with specific campaigns and messages.

## Conclusion

Candidates' experience of a Cambridge ESOL exam is largely through the preparation they receive from their teacher. Their performance in an exam is, in most cases, dependent on the support they receive from that teacher. Cambridge ESOL recognises this and understands how vital it is to provide extensive, relevant and varied teacher support which reaches a worldwide audience. The Teacher Survey reported here highlights the need for further research and development in this area. Changes are already underway, with the development of more classroom preparation packs, an exciting new Teacher Support website offering greater interactivity, bespoke face-to-face training and a central teacher database.

It is clear from the Teacher Survey that teachers value the teacher support Cambridge ESOL provides, but they would like more of it. Cambridge ESOL's goal is therefore to offer a coherent and accessible teacher support package, combining the existing and the new elements to teachers at different stages of their careers.

# Certificating IELTS writing and speaking examiners

**SACHA DEVELLE** RESEARCH AND VALIDATION GROUP, CAMBRIDGE ESOL

## Introduction

IELTS examiners are subject to rigorous procedures for the certification of new examiners and the re-certification of experienced examiners. The IELTS Professional Support Network (PSN) integrates these procedures via a global system of recruitment, induction, training, standardisation, certification and monitoring. Research funded by the IELTS Joint-funded Research Program (Brown 2000, Furneaux and

Rignall 2007), and outcomes from The IELTS Speaking Test Revision Project (1998-2001) and Writing Test Revision Project (2001-2005) have provided valuable insights into examiner training and the regular updating of examiner certification and training materials. Studies in Language Testing 19 (Taylor and Falvey 2007) documents Speaking and Writing studies that have contributed to this ongoing development. Until recently, standardisation and

**Figure 1: The IELTS standardisation and certification process for examiners (adapted from Furneaux and Rignall 2007:426)**



New trainee examiner

Completion of Induction at IELTS centre

Initial face-to-face training: two days with trainer
Studying standardised Writing scripts and Speaking performances at particular band levels

Certification task
Rating standardised Writing scripts and recorded Speaking performances

If successful

Certification: valid for two years

Two years later

Completion of re-certification task

If successful

Certification for a further two years

certification sets were chosen by IELTS Principal Examiners. In 2008 the Research and Validation Group introduced Multi-Faceted measurement (FACETS) as a supplementary quantitative measure to confirm expert judgement. This approach has been used in compiling the latest Cambridge ESOL oral production DVD for use by the Council of Europe (Galaczi and Khalifa 2008, Galaczi and Khalifa 2009). The extension of this mixed methodology to the IELTS context provides the basis of the present article.

## Background to IELTS examiner certification

Qualification as an IELTS examiner starts with a set of Minimum Professional Requirements (MPR). To briefly summarise, potential applicants must be a native speaker or the equivalent of an IELTS Band 9 (an 'Expert User' on the IELTS scale), hold a teaching qualification and have a specified amount of teaching experience. Potential IELTS examiners then attend the induction phase and must successfully complete the following training,

standardisation and certification process. Figure 1 sets out the induction and training cycle for potential and experienced examiners. Of particular interest to the present article is Step 3 which refers to the 'certification task'. The rationale for certification is to assess whether the examiner has understood, and is able to apply, the marking criteria that are used to assess Speaking and Writing performances. Certification is supported by face-to-face standardisation sessions, and standardised Writing and Speaking materials are available to potential and existing examiners during the training cycle. The certification is renewed every two years, as shown in Figure 1.

The standardisation and certification cycles are closely interlinked. Every year a set of standardised Speaking and Writing performances are produced for new and re-certifying examiners. Standardisation sets are produced every two years. These sets provide exemplar ratings for Speaking and Writing performances. A self-access standardisation set is also available at centres, and allows examiners to 'refresh' their knowledge during the re-certification cycle (Taylor 2007:191). The procedure for choosing standardised speaking and writing material involves a multi-phase process that is described next.

## Standardisation/certification set procedures

Each year a large set of potential Writing and Speaking training materials are randomly selected from UK IELTS centres. The process of choosing exemplar samples involves a series of phases set out in Table 1. Phase 1 involves an IELTS Principal Examiner who chooses a pool of representative low, medium and high level performances (Green 2003). Phase 2 involves the remarking of the same Speaking and Writing samples by a group of Principal and Assistant Principal Examiners. Frequency scores and FACETS data from re-marked samples are then produced by the Research and Validation Group (Phase 3) for the data comparison stage (Phase 4). The final decision making process, involving the Research and Validation Group plus Principal Examiners, is carried out in Phase 5. The quantitative FACETS measure, described in Phase 3, is of particular interest to the discussion that follows.

**Table 1: Phase sequence for choice of IELTS standardisation/ certification sets**

| Phase Sequence | Procedure | Decision Makers |
|---|---|---|
| Phase 1 | Selection/filming of Writing and Speaking sample performances | Principal Examiner |
| Phase 2 | Performances re-marked | Principal Examiners Assistant Principal Examiners |
| Phase 3 | Frequency scores FACETS analysis | Research and Validation |
| Phase 4 | Comparison of frequency scores and FACETS data | Research and Validation |
| Phase 5 | Final selection of performances | Principal Examiners Research and Validation |

## Multi-faceted rasch analysis (FACETS)

A particular concern with performance testing (typically Writing and Speaking) has been potential variability in tasks and rater judgments (Bachman, Lynch and Mason 1995). The production of standardised Writing/Speaking sets for certification purposes is no exception. The possible lack of consensus between examiner marks can be assessed by Multi-faceted rasch analysis (FACETS) that provides an account of the harshness/leniency of each examiner and the consistency of each examiner.

FACETS thus allows for the identification of those raters who perform relatively well or poorly and a closer examination of how that rater is performing (i.e., too lenient, too harsh or not consistent). FACETS has been used to examine a range of performance testing issues that include how raters make use of a revised scale (DeVelle 2008a; Vidakovic and Galaczi 2009), experienced and non experienced rater training effects (Weigle 1998) and training/certification of raters (Lynch and McNamara 1996). FACETS data also provides an objective measure for each criterion, offering a level of detail that allows for a comparison of individual criteria, rather than final band score measures that may conceal marking differences (Suto, Greatorex and Nádas 2009:25). The following discussion elaborates on the role of FACETS throughout the performance selection process.

## Score decisions and borderline cases

The Research and Validation Group at Cambridge ESOL provides FACETS data based on examiner, candidate and scale performance. Of particular interest is the 'fair average' score (FAS) that adjusts the 'observed average'[7] to take into account difficulty of the assessment criteria, rater behaviour and candidate ability. Fair average scores are computed for each criterion (i.e., the criteria for IELTS Speaking are Fluency and Coherence, Lexical Resource, Grammatical Range and Accuracy, Pronunciation), and then rounded up or down to match the IELTS band score system. Acceptable scores reflect agreement on all criteria for both frequency counts (computed at band score levels) and the rounded FAS (DeVelle 2008b). Occasionally, borderline scores emerge that are then flagged for closer examination. In such cases Principal Examiners carry out further qualitative analyses of original scores and the written commentary which accompanies the ratings. Performances with two or more borderline criteria are treated as outliers and removed from the data set. The final decision making process, shown in Table 1 (Phase 5) involves collaboration between the Research and Validation Group and Principal Examiners. We now move on to examine how the FACETS programme addresses rater behaviour.

## Rater behaviour

The FACETS summary report also provides statistical indices for rater behaviour. This information allows for a further

reliability check of examiner rating performance. The following data are taken from the production of the 2008 IELTS Speaking Certification set. Of particular interest are the harshness and consistency measures provided for each examiner (N=9). The issue of what is an acceptable range of examiner severity is quite complex and there are no universally accepted rules. Van Moere's (2006) range of −1.00 and +1.00 logits provides a useful cut-off point. Applying these standards, Table 2 shows Examiner 3 rated slightly more harshly (1.33) and Examiner 9 slightly more leniently (−.99). However, these scores are not seriously misfitting on the overall analytic criteria.

Rater consistency levels were compared against Wright and Linacre's (1994) suggested range of 0.6–1.4. With the exception of the borderline score from Examiner 1 (.50), eight of the nine raters fell within the acceptable range, showing that as a group raters performed consistently in their use of the overall scale. It should be noted that all of the Principal and Assistant Principal Examiners who participate in the annual selection of standardised Speaking and Writing performances are experienced raters, and are provided with monitoring feedback on rating behaviour that is not consistent with the group. The use of feedback as a strategy to facilitate accurate judgments is well known in psychological research (Laming 2004) and plays a role in PSN monitoring procedures discussed at the beginning of this article.

## Discussion

To summarise, the present article has described the procedures for choosing standardised examiner materials used during the certification cycle. The use of FACETS in the multi-phase cycle described in Table 1 provides a statistical measure (the rounded FAS) that supplements score decisions for each criterion. Harshness/leniency and consistency estimates of examiner behaviour are also taken into account. It should be emphasised, however, that FACETS cannot measure the cognitive decision-making processes employed by raters, or the criteria they use

**Table 2: Raters' harshness/leniency and consistency**

| Rater | Speaking | All Analytic Criteria | |
|---|---|---|---|
| | **Measure (Logits)** | **Error** | **Infit Mean Square** |
| 1 | 0.24 | 0.23 | **0.50** |
| 2 | 0.24 | 0.23 | 0.51 |
| 3 | **1.33** | 0.23 | 0.93 |
| 4 | −0.12 | 0.23 | 0.90 |
| 5 | −0.38 | 0.23 | 0.861 |
| 6 | −0.33 | 0.23 | 1.36 |
| 7 | −0.17 | 0.23 | 1.16 |
| 8 | 0.18 | 0.23 | 0.94 |
| 9 | **−0.99** | 0.23 | 0.95 |

---

7   The observed average shows the average rating for an element (e.g., if the score is 42 and number of raters is equal to 9 the observed average is 4.67).

individually to rate a performance. This is why cross-referencing rounded FAS scores with frequency data, and subsequent viewing of borderline cases, are necessary qualitative measures during the selection process (DeVelle 2008b, Hubbard and Galaczi 2008). The present work also contributes to monitoring of the IELTS revised pronunciation scale reported in *Research Notes* 34 (DeVelle 2008a). More specifically, rounded FAS scores provide a further quantitative measure of how examiners rate Pronunciation on the full nine-band scale. Operational research, using the combined FACETS indices described here, is currently being undertaken to investigate examiner rating of the revised pronunciation criterion from both performance data and standardised Speaking materials.

## Conclusion

The mixed-method approach described here provides another step towards continuing quality assurance measures necessary for IELTS examiner training. Effective examiner training programmes are essential for the ongoing development of the test, and a key function of the IELTS Professional Support Network (PSN) that provides support, feedback and documentation on the recruitment, induction, training, standardisation, certification and monitoring of examiners worldwide. Twice yearly targeted sample monitoring of practising examiners, and ongoing reliability studies provided by the Research and Validation Group, continue to build on the work presented here.

## References and further reading

Bachman, F, Lynch, B and Mason, M (1995) Investigating variability in tasks and rater judgements in a performance test of foreign language speaking, *Language Testing* 12/2, 238–57.

Brown, A (2000) An investigation of the rating process in the IELTS oral interview in *IELTS Research Reports*, Volume 3, Canberra: IELTS Australia/British Council.

DeVelle, S (2008a) The revised IELTS Pronunciation Scale, *Research Notes 34*, Cambridge: Cambridge ESOL, 36–9.

– (2008b) IELTS *Speaking and Writing certification and standardisation using FACETS*, paper presented at IELTS examiner meeting, Cambridge, 15 September 2008.

Furneaux, C and Rignall, M (2007) The effect of standardisation-training on rater judgements for the IELTS Writing Module in Taylor, L and Falvey, P (Eds) IELTS Collected Papers: Research in speaking and writing assessment, *Studies in Language Testing* volume 19, Cambridge: UCLES/Cambridge University Press, 422–38.

Galaczi, E and Khalifa, H (2008) Exemplifying the CEFR. *The creation and use of a DVD of students' oral language*, paper presented at EAQUALS Conference, Paris, 24 April, 2009.

– (2009) Cambridge ESOL's CEFR DVD of Speaking performances: What's the story? *Research Notes 37*, Cambridge: Cambridge ESOL, 23–9.

Green, T (2003) *Using MFRM to analyse standardisation video performance*,Cambridge ESOL internal report.

Hubbard, C and Galaczi, E (2008) *The life and times of a standardisation DVD for Speaking papers*, paper presented at IELTS examiner meeting, Cambridge, 15 September 2008.

Laming, D (2004) *Human judgment: the eye of the beholder*, London: Thomson.

Lynch, B and McNamara, T (1998) Using G-theory and many-facet Rasch measurement in the development of performance assessments of the ESL speaking skills of immigrants, *Language Testing* 15/2, 158–80.

Shaw, S (2002) The effect of training and standardisation on rater judgment and inter-rater reliability, *Research Notes 8*, Cambridge: Cambridge ESOL, 13–7.

Suto, I, Greatorex, J and Nádas, R (2009) Thinking about making the right mark: Using cognitive strategy research to explore examiner training, *Research Matters: A Cambridge Assessment Publication* 8, Cambridge: Cambridge Assessment, 23–31.

Taylor, L (2007) The impact of the joint-funded research studies on the IELTS Speaking Module in Taylor, L and Falvey, P (Eds) *IELTS Collected Papers: Research in speaking and writing assessment*, Studies in Language Testing volume 19, Cambridge: UCLES/Cambridge University Press, 185–94.

Van Moere, A (2006) Validity evidence in a universal group oral test, *Language Testing* 23, 411–40.

Vidakovic, I and Galaczi, E (2009) ILEC Speaking: revising assessment criteria and scales, *Research Notes 35*, 29–34.

Weigle, S (1998) Using FACETS to model rater training effects *Language Testing* 15, 263–87.

Wright, B D and Linacre, J M (1994) Reasonable mean-square fit values, *Rasch Measurement Transactions* 8/3, 370.

# Using the CEFR to inform assessment criteria development for Online BULATS speaking and writing

**LUCY CHAMBERS** RESEARCH AND VALIDATION GROUP, CAMBRIDGE ESOL

## Introduction

The Common European Framework of Reference (CEFR) 'provides a basis for the mutual recognition of language qualifications, thus facilitating educational and occupational mobility' (COE 2009a) thus for learners there are huge benefits in knowing how their qualifications relate to this framework. The relationship between Cambridge ESOL and the CEFR and details of how the CEFR is embedded in the test development cycle have been well documented (e.g. Taylor and Jones 2006, Cambridge ESOL

2009). The current emphasis is on brining further explicit CEFR reference into Cambridge ESOL's exam processing and documentation. This paper illustrates an example of how this can be built into one of the early stages of the test development cycle; the development of performance assessment criteria.

The Business Language Testing Service (BULATS) forms an integral part of Cambridge ESOL's educational mission, supporting effective language learning in business contexts as a key tool for social and personal development. BULATS is specifically for the use of companies and organisations which need a reliable way of assessing the language ability of groups of employees or trainees. The tests aim to be relevant to people using the language at work; covering areas such as descriptions of jobs, companies and products, travel, management, marketing and customer service. The tasks in the test are generally practical ones, e.g. taking a phone message, checking a letter, giving a presentation, understanding an article, writing a report. BULATS is recognised by companies, labour ministries, education authorities and the tertiary sector globally for recruitment, promotion, developing internal language training programmes and assessing vocational language competencies.

Currently the suite comprises of a Reading and Listening paper (online, CD Rom and paper-based formats), a Speaking test (face-to-face format) and a Writing test (paper-based format). Each test can be used independently of the others, or they can be used in various combinations. The tests are available in English, French, German and Spanish. The tests are all multi-level and assess candidates across the Council of Europe Framework levels A1–C2.

## Development of assessment criteria

Cambridge ESOL is developing online Speaking and Writing tests to further enhance the BULATS suite. These new tests assess candidates' spoken and written production and interaction on a number of dimensions: task achievement, language resource and text organisation for writing, and task achievement, coherence/discourse management, language resource, pronunciation and hesitation/extent for speaking.

One important aspect of the development phase of the test development cycle, is the creation of new assessment criteria which enable candidate performance across these dimensions to be measured. All new/revised performance scales go through a thorough validation process to ensure the scales are measuring accurately and reliably. With the development of the new tests it was decided to extend their measurement range to include pre-A1. New scales covering the full range of the CEFR were written by an experienced consultant who has worked extensively with the CEFR. The next stage was for the individual descriptors within each band to be mapped to the CEFR assessment scales to ensure the new scales accurately reflect the CEFR.

A second consultant, not involved with writing the criteria, was engaged to assess how the new assessment criteria aligned with the CEFR scales. The consultant started by using the Global Oral Assessment Scale (Council of Europe 2009b:184), the Qualitative Aspects of Spoken Language Use (Council of Europe 2001:28–9) and the Written Assessment Criteria Grid (Council of Europe 2009b:187). She took each phrase within the BULATS descriptors for each level and skill and found the CEFR descriptor that was the most similar in content and meaning.

These grids were used for the majority of the 'mappings', however for certain criteria the focus needed to move to more relevant subscales. For example, Pronunciation is an important criteria in the BULATS speaking test but it is not covered in the aforementioned grids; instead the Phonological Control subscale (Council of Europe 2001:117) was used. Likewise, for Online BULATS Speaking the 'business' nature of the tasks pointed towards the use of the Formal Discussion and Meetings subscale for spoken interaction (Council of Europe 2001:78), this was especially so for the coherence criteria. Table 1 illustrates some examples of the mappings at different CEFR levels.

Once the mapping was complete it allowed identification of which descriptors mapped to the CEFR scales and whether they mapped at the correct level. All the descriptors that aligned to the CEFR scales did so at the correct level. However it was found that BULATS descriptors relating to task achievement were not readily matched to the CEFR. This is unsurprising as the CEFR scales focus on

**Table 1: Example draft BULATS descriptors mapped to CEFR assessment criteria (adapted from O'Dell 2009a and b)**

| Phrase from draft BULATS scale | Phrase from CEFR scale | Skill | CEFR level | Which CEFR table |
|---|---|---|---|---|
| A very limited range of language | Very basic repertoire of words and simple phrases | W | A1 | Written Assessment Criteria Grid |
| Range is adequate for some very simple, familiar topics but inadequate for wider topics | Uses basic sentence patterns, with memorised phrases, groups of a few words and formulae in order to communicate limited information mainly in everyday situations | W | A2 | Written Assessment Criteria Grid |
| Pronunciation can generally be understood but L1 features may cause strain | Pronunciation is clearly intelligible even if a foreign accent is sometimes evident and mispronunciations occur | S | B1 | Phonological Control |
| Grammar and vocabulary …. is sufficiently accurate to deal with the tasks | Shows a relatively high degree of grammatical control. Does not make errors which cause misunderstanding | S | B2 | Qualitative Aspects of Spoken Language Use |
| Occasional inaccuracies of grammar or vocabulary may occur | Occasional errors of grammar, collocation or idiom | W | C1 | Written Assessment Criteria Grid |
| Able to express both simple and complex ideas with ease | Can hold his/her own, putting an articulate and persuasive argument | S | C2 | Formal Discussion and Meetings |

performance whereas descriptors of this kind are related to task fulfilment and assessment.

The next stage was to revisit the draft BULATS assessment criteria to see whether they could be improved in light of this exercise. Instances where the descriptors could be fine-tuned so that language used within the descriptor more closely reflected the language of the CEFR grids were highlighted and this will feed into the final revisions of the criteria. These are, for the most part, a word change/addition.

The process described above is only one small piece of the larger alignment picture. Further work is planned on the specification, standardisation and validation of the new tests. However, it does illustrate how the CEFR can be brought into sub-stages of the test development process, being both a useful resource and a way of strengthening alignment. The building of this relationship between Online BULATS Speaking and Writing and the CEFR allows learners' achievements to be recognised within Europe and increasingly beyond as the CEFR becomes accepted more and more as a standard reference of language abilities worldwide.

## References and further reading

Cambridge ESOL (2009) website http://www.cambridgeesol.org/what-we-do/research/cefr/testing-development-cycle.html

Council of Europe (2001) *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*, Cambridge: Cambridge University Press.

– (2009a) website http://www.coe.int/t/dg4/linguistic/CADRE_EN.asp

– (2009b) *Relating language examinations to the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR). A Manual*, retrieved from: http://www.coe.int/t/dg4/linguistic/Manuel1_EN.asp#TopOfPage

O'Dell, F (2009a) *BULATS Writing and the CEFR*, Cambridge ESOL Internal report.

– (2009b) *BULATS Speaking and the CEFR*, Cambridge ESOL Internal report.

Taylor, L and Jones, N (2006) Cambridge ESOL exams and the Common European Framework of Reference (CEFR), *Research Notes* 24, 2–5, Cambridge, Cambridge ESOL.

# CB BULATS: Examining the reliability of a computer-based test

**LAURA COPE** RESEARCH AND VALIDATION GROUP, CAMBRIDGE ESOL

## Introduction

The reliability of a set of test scores is an indication of the amount of measurement error associated with the scores. This level of consistency can be estimated by using methods of internal analysis to compute a reliability coefficient. A coefficient of 0.80 or more for standardised tests generally indicates that the data are reliable enough for practical purposes.

The Business Language Testing Service (BULATS) is a multilingual system assessing four language skills in work contexts. Writing and Speaking are administered individually via the BULATS Writing Test and the BULATS Speaking Test. Candidates have the option of taking either the BULATS Standard Test or the BULATS Computer Test to be assessed on the skills of Reading and Language Knowledge (RLK) and Listening. The computer-based (CB) test is a *Computer Adaptive Test* (CAT) where the computer selects items for candidates according to their responses to the previous items, thereby adapting the difficulty of the items to the level of the candidate. It is divided into two sections, where RLK is tested in the first and Listening is tested in the second section, and candidates receive standardised scores for each section and an overall score for the test.

This article examines the reliability of CB BULATS Version 6.1 using a *Rasch reliability* estimate (an internal consistency measure, analogous to Cronbach's Alpha). The research questions addressed in this study are:

- How reliable is CB BULATS Version 6.1 in terms of the Overall Test reliability and the reliability per section?

- Are the reliabilities found by section and overall satisfactory?

- To what extent does test length affect the reliability of the test?

## Data collection

The data for 1407 candidates who took the English variant of CB BULATS Version 6.1 between March 2006 and June 2006 at various centres worldwide were analysed in this study. These candidates were selected for analysis based on the fact that complete data across both sections of the test was available. It happens from time-to-time that candidates fail to finish the test within the specified time limit of 75 minutes and hence are not awarded a Listening score or an Overall Test score. Such candidates have been removed from the data. Table 1 shows how the candidates varied according to their first language (L1). Unfortunately in CB BULATS candidates are not obliged to answer all questions relating to their background. Hence, it was found that 13% of candidates chose not to give their L1.

**Table 1: Test takers grouped by first language**

| First language | Frequency | Percentage |
|---|---|---|
| German | 407 | 28.93% |
| Portuguese | 278 | 19.76% |
| Italian | 265 | 18.83% |
| Spanish | 135 | 9.59% |
| Russian | 34 | 2.42% |
| Arabic | 18 | 1.28% |
| Turkish | 16 | 1.14% |
| Greek | 8 | 0.57% |
| Polish | 8 | 0.57% |
| Chinese | 6 | 0.43% |
| Farsi | 6 | 0.43% |
| Ukrainian | 6 | 0.43% |
| Other | 35 | 2.49% |
| Not given | 185 | 13.15% |

It can be seen that there is no single L1 group dominating the sample of candidates, however the majority of candidates fall into the Indo-European language group, with the most common languages being German (28.9%), Portuguese (19.8%), Italian (18.8%) and Spanish (9.6%).

The CB response data and test results for the Overall Test and its sub-sections (RLK and Listening) were available in *logfiles* (Microsoft Access Databases with a specific internal structure). Data accumulates within these logfiles in encrypted format automatically upon completion of candidate tests and these files are stored on computer systems at test centres. These logfiles were returned to Cambridge ESOL from the test centres and were decrypted before being assembled into a single database ready for analysis.

## Methodology

After the completion of a CB BULATS candidate test, an estimate of the candidate's ability (measured in logits) is derived from a latent trait Rasch analysis. This estimate is then converted into a BULATS standardised test score (with a range from 0 to 100) by means of a scaling procedure. During the course of this article, the term 'test score' will refer to BULATS standardised test scores and the term 'ability level' or 'ability score' will refer to the candidate's ability as estimated by the Rasch model.

It should be noted that all items administered to the candidates were selected from identical item banks with calibrated item difficulties.

### Estimating the reliability of CB BULATS Version 6.1

The theory of classical reliability proposes that the

observed score is the sum of the true score and an error score, as illustrated by Equation 1.

*Equation 1: Observed scores as a function of true scores*

$X = T + E$

where:
X is the observed score
T is the true score
E is the error score

Statistically, it follows that the variance of the observed scores for a set of candidates is equal to the sum of two separate and uncorrelated variances: one due to the true scores and the other due to the errors of measurement, as shown in Equation 2.

*Equation 2: Observed score variance as a function of true score variance*

$\sigma_X^2 = \sigma_T^2 + \sigma_E^2$

where:
$\sigma_X^2$ is the observed score variance
$\sigma_T^2$ is the true score variance
$\sigma_E^2$ is the error score variance

The Rasch reliability is defined to be the ratio of true score variance to observed score variance, which can be rearranged as shown in Equation 3, to give the reliability in terms of the error score variance and the observed score variance.

*Equation 3: Rasch reliability*

$$RELIABILITY = \frac{\sigma_T^2}{\sigma_X^2} = \frac{\sigma_X^2 - \sigma_E^2}{\sigma_X^2} = 1 - \frac{\sigma_E^2}{\sigma_X^2} = 1 - \frac{RMSE^2}{SD^2}$$

where:
$\sigma_X^2$ is the observed score variance
$\sigma_T^2$ is the true score variance
$\sigma_E^2$ is the error score variance
RMSE is the root-mean-square standard error
SD is the standard deviation of the observed scores

In this study, the standard deviation of the observed scores is simply the standard deviation of the ability estimates of the candidates (SD) and the standard deviation of the errors of measurement is the mean of the associated *standard error of measurement* (SEM) values for the candidates, and is also known as the *root-mean-square standard error* (RMSE).

The ability estimates and corresponding SEM values for the candidates for the Overall Test, the RLK section and the Listening section were extracted from the logfiles and stored in an Excel file in order to estimate the ability standard deviations and means of the SEM values.

Consequently, the Rasch reliability was estimated using Equation 3 for the Overall Test ability scores, the RLK ability scores and the Listening ability scores.

### Estimating the effect of test length on reliability

In order to assess the effect of test length on the reliability of the test, the candidates were divided into two groups: those who had been administered 57 items or fewer and those who had been administered 58 items or more. This cut-off ensured that the numbers of candidates within each

group was as equal as possible. Following this division, the reliability was calculated for each group using the previously described method.

The relationship between reliability and test length theoretically follows a mathematical relationship, which is described by the Spearman-Brown formula in Equation 4. This relationship requires that both groups of examinees are statistically equivalent in terms of their abilities on the construct being tested, that equivalent item banks are used and that the conditions under which the test is administered are consistent between the groups.

*Equation 4: Spearman-Brown formula*

$$\rho_n = \frac{n\rho_0}{1 + (n-1)\rho_0}$$

where:

$\rho_n$ is the reliability of the lengthened test

$\rho_0$ is the reliability of the original test

$n$ is the factor by which the test is lengthened

The mean number of items administered was calculated for each group so that the factor by which the test was lengthened could be derived. The theoretical predicted difference in reliability and observed difference in reliability between the two groups were both calculated and compared.

## Findings and discussion

### Reliability of CB BULATS Version 6.1

Table 2 gives some statistical characteristics of the candidates' test data, namely the average length (in items) and the standard deviation of the ability estimates for the overall test and its two sub-sections: RLK and Listening.

**Table 2: Statistical characteristics of the test**

|  | Overall Test | RLK section | Listening section |
|---|---|---|---|
| Mean test length (items) | 57.3 | 32.4 | 24.9 |
| Ability standard deviation (SD) | 1.22 | 1.21 | 1.61 |

As can be seen in Table 2, the RLK section is noticeably longer in length than the Listening section, which is in-line with the paper-based version of BULATS. Also, the Listening section has a higher standard deviation of abilities than the RLK section.

The Rasch reliability estimates for the Overall Test, the RLK section and the Listening section are shown in Table 3.

**Table 3: Rasch reliabilities for the Overall Test, RLK and Listening sections**

|  | Overall Test | RLK section | Listening section |
|---|---|---|---|
| Ability standard deviation (SD) | 1.22 | 1.21 | 1.61 |
| Mean SEM (RMSE) | 0.30 | 0.40 | 0.45 |
| Rasch reliability | 0.94 | 0.89 | 0.92 |

The reliability of the Overall Test is 0.94, which is very high. At first glance, it is also surprising that the reliability of the Listening section (0.92) is higher than that for the RLK section (0.89) since the Listening section is significantly shorter than the RLK section. However, the Listening section had a higher standard deviation of ability estimates which would have helped to raise this reliability. Reliability improves as the collection of score data becomes more widely spread from the mean and the range of scores increases.

### Effect of test length on reliability

Table 4 shows the statistical properties of the Overall Test for two sub-groups of examinees: those who had been

**Table 4: Statistical characteristics for each sub-group of candidates**

|  | Group answered 57 items or fewer | Group answered 58 items or more |
|---|---|---|
| Number of examinees | 673 | 734 |
| Average length (items) | 54.98 | 59.49 |
| Mean ability | 0.33 | 0.44 |
| Ability standard deviation (SD) | 1.20 | 1.23 |

administered 57 items or fewer and those who had been administered 58 items or more.

As can be seen from the table, there are comparable numbers of examinees within each group, the candidates' mean abilities were close and the standard deviations of their abilities were very close. It is therefore safe to assume that the two groups of examinees are statistically very similar in the way that they have performed on CB BULATS Version 6.1. Also, there was no distinction in the way that the test was administered for each group or in any other factor relating to the group such as motivation, instruction etc. Hence the Spearman-Brown relationship can be applied in this situation.

Table 5 shows the calculation of the Overall Test reliability for these two groups of examinees.

**Table 5: Overall Test Rasch reliability for each sub-group of candidates**

|  | Group answered 57 items or fewer | Group answered 58 items or more |
|---|---|---|
| Ability standard deviation (SD) | 1.20 | 1.23 |
| Mean SEM (RMSE) | 0.30 | 0.30 |
| Rasch reliability | 0.938 | 0.941 |

The figures for the reliability were very close (0.938 for the group who were administered shorter tests and 0.941 for the group who were administered longer tests), however the higher reliability was produced for the group of candidates who took the longer tests, as would be expected.

Using the Spearman-Brown formula, the predicted reliability of the group of examinees who were administered

a longer test can be calculated. The ratio, *n*, of the average test lengths is 1.08 (calculated using the data in Table 4). The reliability of the group with the shorter average test length was 0.938. Hence, the predicted reliability of the group with the longer average test length is 0.943. This is slightly higher than the observed value (0.941) which means that the reliability of the test is not influenced quite as strongly by the test length as would be expected.

So it would seem that the length of the test in CB BULATS Version 6.1 has less of an impact on reliability than predicted. One possible explanation for this is that candidates with extreme abilities (lower than −3.3 logits or greater than 4.4 logits) tend to have very short tests. The test will stop automatically (provided that a minimum number of items have been administered) if a candidate's estimated ability moves outside of these limits. Hence, the extreme ability candidates will fall into the shorter test length group, creating a higher standard deviation of abilities and consequently a higher reliability.

## Conclusion

The main objective of this study was to examine the reliability of CB BULATS Version 6.1 in order to produce a reliability figure for the test. This is an important consideration because a high reliability is an indication that a test is accurate in the rank-ordering of candiates and consistent in the scoring of candidates across repeated administrations of the test.

It has been demonstrated that the reliability coefficient for the Overall Test is 0.94. It has also emerged that the

reliabilities of the test's sub-sections (RLK and Listening) are high, coming out at 0.89 and 0.92 respectively. As a result, we can claim that the reliabilities of CB BULATS Version 6.1 and each of its two sub-sections are suitably very high.

It has also been shown that the effect of the test length on reliability is lower than the effect predicted using the Spearman-Brown relationship, indicating that candidate ability estimates are not necessarily less reliable if fewer items are administered. This is a positive finding since the nature of the CAT tests means that candidates will have differing lengths of tests, for various reasons such as the estimated SEM falling below a set threshold, however it has been shown that this does not impact on the reliability of the ability estimates.

This study forms part of ongoing research into the reliability of CB BULATS. It would be interesting to investigate how the test scores for CB BULATS Version 6.1 are influenced by test-taker features such as gender, age, familiarity with computers, country of origin etc.

## References and further reading

Geranpayeh, A (2001) *CB BULATS: Examining the reliability of a computer-based test using test-retest method*, Cambridge ESOL internal report.

Jones, N (2000) BULATS: A case study comparing computer based and paper-and-pencil tests, *Research Notes* 3, 10–3, Cambridge: Cambridge ESOL.

Traub, R E and Rowley, G L (1991) NCME Instructional Module: Understanding Reliability, *Educational Measurement: Issues and Practice* 10/1, 37–45.

# Conference reports

The Research and Validation Group have participated in a range of national and international and events recently, presenting workshops and papers on a range of topics, reported on below by Fiona Barker, Guy Nicholson, Martin Nuttall, Szilvia Papp and Angeliki Salamoura.

## 36th ALTE Meeting and Conference, Santiago de Compostela

The 36th ALTE Meeting and Conference took place in Santiago de Compostela, Galicia, Spain from 22–24 April 2009 at the San Francisco Hotel and was attended by a number of colleagues from Cambridge ESOL. Angeliki Salamoura (Cambridge ESOL) and Cecilie Carlsen (University of Bergen, Norway) led a workshop entitled *Linking learner corpora to the CEFR* in which they presented two projects, firstly, the linking of the Cambridge English Profile Corpus and, secondly, the linking of ASK (the Norwegian Learner Corpus) to the CEFR. The current lack of well-defined proficiency levels in many learner corpora mean that corpora are not yet used as fully as they might be within language testing and assessment. This workshop

aimed to show that linking learner corpora to the CEFR would make them a more useful tool for language testers and would make an empirical validation of the CEFR level descriptors possible. Workshop attendees participated by placing learner texts from these corpora on to different CEFR scales.

Other workshops included one by Shelagh Rixon from the University of Warwick, and one by Alan Davies, Emeritus Professor at the University of Edinburgh. Shelagh's workshop was entitled *Making large-scale tests of language for young learners both feasible and child-friendly* in which she looked at what aspects of language use are appropriate to try to assess in children, what test item types children respond to well and less well, and how large-scale testing can reconcile the needs of children with the need for practical test feasibility. Alan led a workshop on *Issues involved in testing academic language* in which he brought together a panel of representatives from six testing bodies that offer tests of academic language in the following languages – Bulgarian, Danish, Dutch, English, French, and German – to discuss the issues and challenges involved in testing academic language. Cambridge ESOL was

represented on the panel by Dr Nick Saville, Director of Research and Validation, who talked about the issues in relation to IELTS.

ALTE was honoured to have the attendance of Leonard Orban, EU Commissioner for Multilingualism, on the conference day. The Commissioner gave the opening address and talked about *Promoting intercultural dialogue, linguistic diversity and language learning*. Plenary sessions were then given by Shelagh Rixon on *Feasibility, Child-Friendliness and Positive Washback – a tricky trio for testing Young Learners*, Alan Davies on *Changes in academic English proficiency assessment in the UK: an explanation*, David Bearfield, Director, European Personnel Selection Office on *Recruiting for the Tower of Babel*, Alison Graves, Senior Interpreter, Directorate General for Interpretation and Conferences, European Parliament on *Multilingualism in Action: the European Parliament*, Professor Barry O'Sullivan on *Standard Setting, the CEFR and tests of Language for Specific Purposes*, and Margarita Chamorro on *The Galician assessment system: the CELGA certificate*s.

Around 120 people, including many local participants, attended the meeting and conference, and some twenty-five people from several countries stayed on to attend a two-day introductory course on Testing Language for Specific Purposes (LSP) on 25–26 April which was taught by Professor Barry O'Sullivan from the University of Roehampton, UK. The course focused on three key issues in testing LSP – identifying and reporting success, content and language boundaries, and establishing evidence of specificity.

## EALTA 2009, Turku

The 6th Annual EALTA Conference took place in the Finnish city of Turku in the first week of June 2009. Attended by over 200 speakers and delegates from around Europe, the conference was held over four days at the Department of Teacher Education at the University of Turku.

The theme of this year's conference was *Synergies and Tensions in Language Testing and Assessment*. Among the keynote speakers was Lyle Bachman from the University of California who, in a reprise of a talk given to staff at Cambridge ESOL in April 2009, spoke about his development of the *assessment use argument* as a conceptual framework, to make language assessors more accountable to all stakeholders in the assessment process by providing a rationale and justification for all decisions made in the design and development of language tests. Another keynote paper was presented by Christer Laurén from the University of Vaasa, who discussed the impact of total language immersion programmes in Canada, Spain and Finland.

Numerous individual papers were presented at the conference which related to various current issues affecting language assessment in Europe. Among these were a paper on the perceived impact of the new *Nowa Matura* English school-leaving exam in Poland, given by Jo Lewkowicz from the University of Warsaw and Elzbieta Zawadowska-Kittel from the Higher School of Linguistics, and a paper on the development and delivery of new national computerised English language tests for primary and secondary school

students in Norway, given by Anders Fikke Johannesen and Hildegun Lahlum Helnes from the University of Bergen.

Another well-received paper was presented by Anthony Green from the University of Bedfordshire, who spoke about his research study (with Roger Hawkey of the same institution) into the selection and editing of texts and tasks for the IELTS Academic Reading test, and the differing approaches employed by items writers in this regard. Cambridge ESOL was also represented at the conference, with Evelina Galaczi from the Research and Validation Group presenting a paper on the development of the revised rating scales for ESOL Main Suite Speaking tests.

The conference concluded with a panel discussion convened by Christine Niakaris from the Hellenic American University, which reviewed many of the issues raised over the course of the preceding four days. Among those debating these issues were Neus Figueras from the Departament d'Educació, Generalitat de Catalunya, Tommy Lagergren from the Swedish National Agency for Education, Jo Lewkowicz from the University of Warsaw and Lyle Bachman from the University of California.

## SLATE meeting, Jyväskylä

Angeliki Salamoura attended the latest SLATE Meeting which was hosted on 10 June 2009 in Jyväskylä, Finland, by the Department of Languages and the Centre for Applied Language Studies of the University of Jyväskylä. SLATE (Second Language Acquisition and Testing in Europe) is a European research network that combines expertise in language acquisition and assessment to study the Common European Framework of Reference for Languages (CEFR). The network has organised two international workshops to present CEFR-related research and holds biannual meetings to plan further research and cooperation. The English Profile Programme is linked with the network and shares its aims. One of the main issues discussed in the June 2009 Meeting was the planned publication of an edited volume presenting the research work of SLATE-related projects. English Profile will contribute a chapter in this volume which will illustrate how hypotheses formulated from theories and models of second language acquisition (SLA) and psycholinguistics, and a corpus-informed approach have been used to investigate second language learner data in order to develop the Reference Level Descriptions for English.

## Learning and Assessment in Primary Schools Conference, Cambridge

The English for Schools 'Learning and Assessment in Primary Schools' conference took place in Cambridge between 12–14 June 2009. It is the second international event held in Cambridge after the first seminar in 2004 and eighty delegates from 22 countries attended the conference. Two keynote speakers were Annie Hughes from the University of York and Lyle Bachman from University of California, Los Angeles. From Cambridge ESOL, Szilvia Papp talked about the development of Can Do statements for school learners. She presented the research and validation

activities that led to the development of Can Do statements which now appear on KET for Schools and PET for Schools certificates and handbooks. She discussed with delegates the way Can Do statements can be used in formative assessment in the classroom and explained that the development of a comprehensive list of Can Do statements to be used with younger learners is an ongoing activity for Cambridge ESOL. Neil Jones introduced the European Survey of Language Competences as an example of the use of the Common European Framework of Reference as an instrument of language policy making in Europe. He pointed to the need to link such external objectives to practical classroom work, and presented Asset Languages as an example of an attempt to link formative teacher assessment into a proficiency framework. At the end of his talk, he presented parts of a possible toolkit that examination boards such as Cambridge ESOL can develop to help teachers in assessing their learners formatively. All presentations given at the conference can be viewed at http://cambridgeesol.gdbtv.com/

## ALTE Language Testing Courses 2009

Two interesting and stimulating courses on language testing were run at Laboratorio Itals, Universita Ca' Foscari in Venice by ALTE, in September, led by Cambridge ESOL and external experts. The first course took place from 7–11 September and was an Introductory Course in Language Testing. The course was taught by Dr Lynda Taylor (Consultant to Cambridge ESOL) and Professor Cyril Weir (University of Bedfordshire). Twenty-five participants from some eleven countries attended sessions focussing on the practical application of testing and assessment theory. These included topics such as understanding the test production process, writing test materials, managing test delivery, and building the test validity argument.

The second course, on Testing Listening, took place from 14–18 September at the same location and was led by Dr John Field (University of Reading) and Dr Ardeshir Geranpayeh (Cambridge ESOL). Twenty-seven participants from twelve countries attended sessions on a range of topics including listening test design and production, listening test delivery, and the validation of listening tests.

ALTE also ran an introductory course on Assessing Speaking just before the ALTE Meeting and Conference in Maynooth, Ireland in November 2009. The two-day course took place on 9–10 November and was taught by Dr Lynda Taylor, Consultant to Cambridge ESOL. The course covered a range of topics including speaking test design and production, speaking test management and processing, and the validation of speaking tests.

Further testing courses will take place in 2010 and full details of these and other ALTE activities can be found on the ALTE website: www.alte.org

# Cambridge ESOL PhD Scholarships awarded for English Profile research

Cambridge ESOL, part of Cambridge Assessment, has awarded two PhD scholarships to contribute to the English Profile Programme - a unique interdisciplinary research programme that will provide the first complete set of reference level descriptors of the English language (see www.EnglishProfile.org). Dr Nick Saville, Director of Cambridge ESOL's Research and Validation Group, states:

'English Profile's core objective is to develop a research-based and verifiable way of understanding exactly how people learn English, and what they learn at each stage of that process. The value of this for learning, teaching and assessment will be enormous. Cambridge ESOL is proud to support posts which will advance this exciting second stage of English Profile.'

The two successful applicants commenced their studies in October 2009 at two departments of the University of Cambridge which are involved in English Profile research. Stephen Spencer has been admitted as a PhD student at the Research Centre for English and Applied Linguistics (RCEAL) working on Computational Linguistics under the supervision of Dr Paula Buttery. Stephen holds a BSc in Mathematics from the University of Warwick and an MSc in Computing from Imperial College London. He held the position of research scientist at QinetiQ, a UK defence technology company, for five years, working in the Data Mining team specialising in machine learning and natural language processing. Stephen is investigating criterial, discriminative linguistic features of the different Common European Framework of Reference (CEFR) proficiency levels, exemplified in the Cambridge Learner Corpus and other learner texts.

Helen Yannakoudakis also started her PhD programme in October 2009 and is based in the Computer Laboratory, University of Cambridge, working on Computational Linguistics supervised by Professor Ted Briscoe. Helen undertook the MPhil in Computer Speech, Text and Internet Technology at the Computer Laboratory in 2008–9 and her project thesis was entitled *Grapheme-to-Phoneme Conversion for Greek*. Before coming to Cambridge Helen was a student at Athens University of Economics and Business where she specialised in Computer Science. Helen will investigate the use of data/text mining and knowledge discovery techniques as an empirical and systematic method for uncovering criterial, discriminative linguistic features of the different CEFR levels.

We would like to welcome Helen and Stephen to the growing network of English Profile researchers and look forward to reporting on their research in future issues.