# English

**UNIVERSITY** *of* **CAMBRIDGE**
ESOL Examinations

**English for Speakers of Other Languages**

## INDEX

**UNIVERSITY of CAMBRIDGE**
**ESOL Examinations**
**English for Speakers of Other Languages**

# Introduction to UCLES

UCLES is a department of the University of Cambridge. The UCLES Group comprises three business units. Each business unit has a distinct product range and group of customers. Cambridge ESOL (English for Speakers of Other Languages) provides examinations in English as a foreign language and qualifications for language teachers throughout the world. CIE (University of Cambridge International Examinations) provides international school examinations and international vocational awards. OCR (Oxford, Cambridge and RSA Examinations) provides general and vocational qualifications to schools, colleges, employers, and training providers in the UK. As a major examination board, UCLES plays an important role within societal and educational processes by providing examinations for a wide range of purposes and educational contexts, both in Britain and in around 150 countries worldwide.
See: www.ucles.org.uk

# Cambridge ESOL

Cambridge ESOL's main suite of certificated examinations has a long history dating back to 1913 and over the years new examinations have been introduced and existing exams have been revised on a regular basis. The first examination to be developed was the CPE (first introduced in 1913) and it was most recently revised in the late 1990s (with the revised examination launched in December 2002). This is the C2 level in the Cambridge main suite. The most widely taken examination, FCE was first introduced in 1939 and was most recently revised in 1996. During the latter half of the 20th century this test became very well known amongst EFL teachers and the ELT industry and was widely accepted as a common standard at intermediate level. The authors of the CEFR used this level as one of the starting points for the development of the CEFR scales.  In the Cambridge main suite FCE is at B2 level. CAE (C1) was introduced in 1991, KET (A2) in 1994, and the revised PET (B1) in 1995.  In 2004/5 the harmonised suite of examinations constitutes a five level system of criterion-related examinations. This level system has been linked to the framework of the Association of Language Testers in Europe (ALTE - 1991) and subsequently to the Council of Europe's Common European Framework of Reference (2002) ranging from A2 to C1 on the CEFR scale.

The ESOL scale provides a set of common standards and is the basis of the criterion-referenced approach to the interpretation of examination results. Referencing to the criterion is undertaken by means of scalar analyses using the Rasch model to relate the results from the whole range of Cambridge examinations to a Common Scale.  The production and validation of test materials and assessment procedures through Quality Management Systems ensures that numerous quality checks are implemented at all stages in the process - from commissioning new test material to examinations being administered at testing centres around the world. This helps to ensure that the examination system as a whole is consistent and stable over time.

See: www.CambridgeESOL.org

## Test Development

Cambridge ESOL seeks to achieve the overall usefulness of its examinations by working closely with a full range of stakeholders to ensure good "fitness for purpose". The specifications for each component of an examination provide a clear definition and detailed description of what in being tested (in terms of constructs and test content) and what must be produced for that test by the item writing team. These provide the basis for the detailed item writer guidelines which are used by the Subject Officers and the item writers to ensure that test materials match the underlying constructs and intended content of the test.

For item based examinations of Reading and Listening this process consists of eight main stages: commissioning, pre-editing, editing, pretesting, pretest review, paper construction, examination overview, and question paper production (QPP). Under operational conditions the cycle of production for the examinations usually begins at least 2 years before the examination itself takes place.

Question paper production is based on the Local Item Banking System (LIBS), which is a computer-based management and analysis tool developed by UCLES to handle the entire production cycle. LIBS contains a large bank of materials for use in the examinations (more than100,000 items) which have all been fully edited and pre-tested according to the procedures described below.

## Eight Stages:

### 1 Commissioning

Commissioning of item writers is the first stage of the question paper production process. The item writers, who are all external consultants, are commissioned to produce a specific range of questions based on the specially adapted versions of the Test Specifications (i.e. Item Writer Guidelines). It typically takes 15 weeks from commission to the submission of the questions.

### 2 Pre-editing

Pre-editing takes place when the commissioned materials are received by Cambridge ESOL for the first time. Subject Officers scrutinise the material submitted, to assess its suitability for use as test items, and to reject unsuitable, problematic or weak material. Material is screened for taboo or sensitive topics which are likely to be biased against particular groups of candidates.

### 3 Editing

Materials which successfully pass the pre-editing stage are submitted for editing. At this stage item writers check or recheck the quality of material against Specifications and Item Writer Guidelines and make any changes necessary to submitted materials so that they are of an acceptable standard for pretesting (i.e. appear to be of an acceptable standard for inclusion in a live test). They also ensure that the key, rubric, tapescript, etc. are accurate and comprehensive. All item writers are involved in the editing of their own items in the context of a team led by a Subject Officer and/or a senior external consultant.

## 4 Pretesting

In order to confirm the quality of materials, Cambridge ESOL uses the process of pretesting. Every year, around 30,000 candidates are involved in the pretesting of Cambridge ESOL examinations. Almost all of them are learners who are preparing for or have recently taken a Cambridge ESOL examination. They take the pretest under examination conditions, and their answers are assessed in the same way as 'live' examinations. Each edited item is pretested on a representative sample of candidates (usually involving around 200 learners who are about to take one of the live examinations) so that data can be statistically analysed. In this way pretesting plays an important role in achieving reliability in terms of parallel forms of the tests.

All the materials which are pretested can be related to the underlying scale of difficulty in the item bank by the use of 'anchor' items; these are items with known measurement characteristics which provide the basis for calculating the difficulty of the new items to go into the bank. Cambridge ESOL uses the Rasch model to construct the common scale which underpins the item-bank and which provides the basis for the construction of parallel forms of the tests at the different levels of the system.

## 5 Pretest Review

After pretesting, a meeting is held to review the performance of materials. The meeting reviews the results of the pretests and evaluates the measurement characteristics of tasks and items. Any essential adjustments to tasks and items are made at this stage, ensuring, as far as possible, that no editing will need to take place at the paper construction stage. The meeting also scrutinises the marking keys provided for each item by the Item Writers.

## 6 Paper Construction

At this stage items are taken from the item bank and combined to form complete papers according to established procedures, using the Local Item Banking System as a test construction tool. This allows the construction of complete papers that assess the full range of skills, contexts, etc, using materials that have all been fully pretested. Paper construction normally takes place about two years before the date of the live examination. Material is selected for a first draft of the question paper and recommendations are made to the team which will check that:

- a range of topics/tasks is maintained in the paper according to the Specifications
- there is no obvious overlap in content across the parts in the paper
- an examination paper as a whole possesses the required continuity
- a complete set of statistics and other information is available from the Local Item Banking System.

The common scale which underpins the item-bank, based on Rasch scaling, means that the paper construction can be based on a target for the average difficulty of the paper using the difficulty estimates obtained from pretesting. When a new paper has been constructed, reports can be generated from LIBS which shows the mean difficulty of the paper as a whole and the distribution of item difficulty across the different tasks in the paper. This method provides a prediction of how difficult the paper will be when it is used under live conditions. After the examination has been administered and scored, the grading process provides additional confirmation of the difficulty of the paper and allows for adjustments to be made if necessary in order to maintain a consistent standard across different versions.

## 7 Examination Review

Once all of the papers have been constructed, they are brought together to form a complete examination, where each paper covers the prescribed range of content and skills. This is then considered as a complete examination by Subject Officers and Examinations Managers as part of the examination review.

## 8 Question Paper Production (QPP)

Final copies of the constructed question papers are passed on to a Question Paper Production Unit (QPP) no later than 18 months before live test administration for Main Suite. The papers go through approximately 20 subsequent stages (depending on the examination) in order for the manuscript to be transformed into error-free camera-ready copy (usually in the form of PDF files) and finally into printed question papers which are sent to examination centres. At this stage in the process, several rounds of proof-reading and content checking are used to provide additional checks on the quality of the materials.

## Statistical Analyses

Statistical evidence is important in providing end users with good evidence for the quality and fairness of the exams. All the ESOL test materials are analysed before being used and are stored in the LIBS item bank, linked to an underlying scale using Rasch-based methods. In carrying out post-exam analysis Cambridge ESOL employs both classical methods and Rasch-based methods on a routine basis. The overall grade of a main suite examination is derived from multiple observations across all components. For each component, attempts are made to ensure that the reliability of the assessment is as high as possible. In addition, the dependability of the final grade is checked by the procedures used during the grading and awards procedures. Cambridge main suite examinations report results as a single overall grade (e.g. A to E); the reliability of the examination can be estimated as a composite (as discussed in Feldt and Brennan, 1989, and Crocker and Algina, 1986) The typical composite reliability of the ESOL exams is generally considered acceptable; for FCE, for example, it is estimated at 0.94.

ESOL sets target levels for the internal consistency reliability for the item-based components of the main suite examinations using Cronbach's alpha. These target levels are routinely used in the test construction procedures and the predicted operational reliability for each paper is based on the type and quality of the tasks which are chosen according to the test specifications. The information used includes the Rasch-based difficulty estimates and other data obtained during the item writing and pre-testing processes.