

Quality Assurance and Quality Control: Reviewing and pretesting examination material at Cambridge ESOL

TONY GREEN, RESEARCH AND VALIDATION GROUP
DAVID JAY, PRETESTING UNIT

Introduction

Cambridge ESOL is recognised by universities, employers and national education authorities around the world for its professional, rigorous and high standard assessments. This article will describe the vital role played by the reviewing and pretesting of material in ensuring the quality of our examinations.

Quality Assurance and Quality Control

When considering the effectiveness of an organisation in delivering quality, a distinction is often made between *quality assurance* and *quality control*. Quality assurance is concerned with processes and involves the management of activities and resources to improve benefits to stakeholders. Quality control is concerned with outcomes and involves checking that products meet intended standards. When producing examinations we follow procedures calculated to generate material of high quality, judge this material against established standards for quality control and feed back results to refine our processes.

Figure 1: The Question Paper Production process at Cambridge ESOL

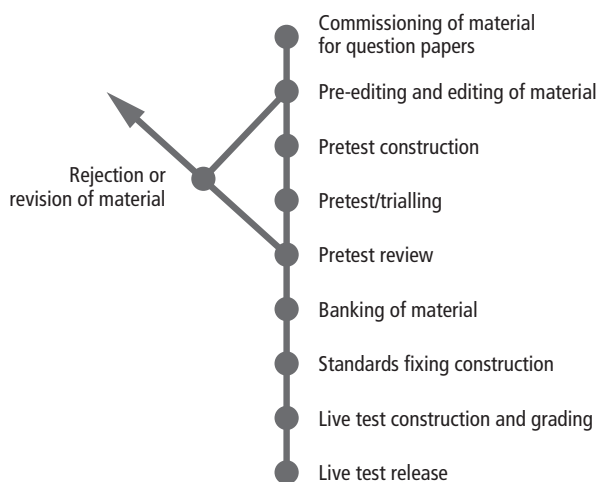


Figure 1 shows in outline the process of question paper production at Cambridge ESOL. At each stage in this process, from the initial commissioning of test material to the assembly of live test papers, there are checks in place to ensure that material

reaching the live tests is of the highest quality. Quality control checks in question paper production take two forms: inspection and testing. Inspection involves people looking at new material to check whether it conforms to specifications; testing involves investigating the material to find out how it performs in use. Key occasions for inspection come at the *Pre-editing* and *Editing* stage and at *Test Construction*. Testing is of particular relevance to the *Pretest Review* and *Test Construction* stages.

Quality Control 1: initial inspection of the test material

Pre-editing is the first stage of the editing process and takes place when commissioned materials are initially submitted by item writers. A meeting is held involving the chairs of the item writer teams (experienced item writers) and Cambridge ESOL staff to consider the material. At this stage, guidance is given to item writers on revising items and altering texts, and feedback is provided on rejected texts and/or unsuitable item types. Routine checks at this point are intended:

- to ensure that all test material is culturally appropriate and accessible world-wide
- to ensure that all test material meets the test specifications
- to suggest appropriate changes to materials requiring amendments or rewriting.

With respect to these considerations, the *Pre-editing* process includes attention to the following features of the material:

- Topic
- Topicality
- Level of language
- Suitability for the task
- Length
- Focus of text
- Style of writing
- Focus of task
- Level of task

At *Editing*, texts and selected items are scrutinised again and are either approved for pretesting, amended, or, occasionally, may be sent back to a writer for further revision. Revised material will then be re-submitted for editing at a subsequent meeting.

Quality Control 2: testing the test material

In addition to the rigorous inspection of new material against test specifications, Cambridge ESOL also investigates the material in use before banking it for test construction. All material destined for use in live tests is first piloted through *Pretesting* (of papers containing items that can be marked objectively by computer or by clerical markers such as Reading papers) or *Trialling* (of papers that are scored by professionally trained examiners such as Speaking papers). New tasks are assembled into test forms and distributed to a sample of candidates preparing for Cambridge ESOL examinations. Candidates participating in pretesting and trialling are broadly representative of the population taking the equivalent live test. Results from Pretesting are analysed by the Research and Validation Group and decisions are then made about whether the material can be banked for use in live tests. The following section describes some of the challenges presented by this process and the procedures in place to manage these effectively.

Quality Assurance: managing the pretest process

To commence the pretesting process, invitations are sent out to centres world-wide, targeting both registered and potential candidates for specific examinations. *Pretest windows* – the periods when pretests are made available to centres – are generally scheduled between 6 to 12 weeks prior to live test dates so that candidates for pretests will be taking the live examination soon afterwards, but will have enough time to benefit from feedback.

Response validity is enhanced by psychological and linguistic readiness on the part of candidates. Inevitably, there are substantive differences between pretest and live administration, not least the fact that the majority of pretests are invigilated by teachers in classrooms to students with varying degrees of commitment to taking a practice examination. Centres are therefore requested to place an emphasis on simulating the live experience in a 'dry run' of the examination day. It is important that the invigilation of the pretest should mirror that of the live examination. This level of authenticity is important, both for Cambridge ESOL and for the candidates themselves.

In the first instance, Cambridge ESOL relies on the good will and the commitment of centres to the quality of our examinations to encourage them to participate in pretesting. A significant source of motivation is that papers are returned to Cambridge ESOL for marking, with raw scores being returned to candidates within 3 weeks of receipt. However, as question paper security is paramount, there are limitations on the amount of feedback we are able to provide to candidates.

Finding enough pretesting candidates for our examinations presents Cambridge ESOL with a significant ongoing challenge, particularly for examinations with a more limited live candidature. To give some idea of the scale of this operation, almost 1,000 different pretest and trial papers were completed during the year 2003–4 with over 1,400 being scheduled for the year 2004–5. In order to meet the needs of examinations with limited candidature, in addition to the existing pretest windows, centres may be encouraged to increase candidature by the offer of pretesting on demand outside the allotted windows, through weekend 'Open Days' at larger centres or through a variety of incentive schemes.

For every live examination constructed, several pretest versions must be processed in order to accommodate the rejection or revision of test items. The average period between a pretest session and the appearance of the material on a live examination paper is 2 years. As a further guarantee of security, each live test paper is made up of tasks drawn from a number of different pretest versions.

The target specification for statistical analysis is a representative sample of 250 candidates per pretest version. In order to reach these numbers, over 300 question papers are allocated for each version and despatched world-wide. A further sample specification requires that there should be no more than 30% of speakers of any first language in a data set. All participants are volunteers and requests for numbers of papers sent to Cambridge ESOL by centres can only be estimates, so it is inevitable that rates of return are sometimes lower than expected. On average, just over 60% of papers are sent back completed. The returned sample may then be further pared down once item responses have been marked and optically scanned. Files passed on for statistical analysis will contain only those candidates who have completed the pretest correctly, have been entered at the appropriate level, and match the targeted balance of first language groups. Thus a proportion of completed pretests may be removed prior to statistical analysis. If the number of returned pretests is too low to meet the requirements material is re-allocated for use in the following pretest window.

In addition to the pretests, all candidates are also administered an anchor test and a background questionnaire. The anchor tests are made up of items of known difficulty, and this allows us to estimate the difficulty of each pretest version in relation to the established Cambridge ESOL Common Scale. Candidate background data (including, among other variables, age, gender and first language) enables Cambridge ESOL staff to further investigate, and if necessary, modify the sample for each pretest. Version-specific questionnaires for invigilators and candidates are also included and provide an invaluable source of qualitative feedback in evaluating the test material from the users' perspective.

Analysis of pretest results

At this point data files that satisfy requirements for representativeness are forwarded to Research and Validation for analysis. Objectively marked papers (papers that are marked clerically or by computer) are treated differently to subjectively marked papers (papers marked by qualified language teachers trained as examiners).

Objectively marked papers: Listening, Reading and Use of English

All candidate responses are analysed to ascertain the measurement characteristics of the material and match these against established standards. Both classical item statistics and latent trait (Rasch) models are used in order to evaluate the effectiveness of the test material¹. Classical item statistics are used to identify the performance of a particular pretest in terms of the facility and discrimination of the items in relation to the sample that was used. Rasch analysis based on the anchor test is used to locate items on

1. For more on the Classical and Rasch approaches see Bachman 2004.

the Cambridge ESOL Common Scale of difficulty. In addition, comments on the material by the pretest centres and the immediate response of the pretest candidates are collected and taken into account in the selection process.

Subjectively marked papers: Writing and Speaking

Writing pretest scripts are marked by senior examiners and their comments are scrutinised to assess the suitability of tasks for inclusion in live test versions. The feedback on the trialling of the Speaking tasks is also assessed before material can be banked.

At a *Pretest Review* meeting, the statistics, feedback from candidates and invigilators and any additional information are reviewed in relation to Cambridge ESOL standards and informed decisions are made on whether texts and items can be accepted for construction into potential live versions. Material is then banked to await test construction.

Banking of material

Cambridge ESOL has developed its own item banking software for managing the development of new live tests. Each section or task is banked with statistical information as well as comprehensive content description. This information is used to ensure that the tests that are constructed have the required content coverage and the appropriate level of difficulty.

At regular *Test Construction* meetings, test papers are compiled according to established principles. Factors taken into account include:

- the difficulty of versions and the spread of items across defined ranges
- the balance of topic and genre
- the balance of gender and accent in the Listening versions
- the range of skills tested.

Our item banking software allows the test constructor to model various scenarios in order to determine which tasks should be combined to create tests that best meet the specifications.

Conclusion

As Cambridge ESOL continues to expand its range of examinations, generating, reviewing and pretesting sufficient material becomes ever more challenging. To safeguard quality we must continue to develop our quality assurance procedures and find new and more efficient means of inspecting and testing our material while enhancing their rigour. For the future, computer and internet-based solutions may provide promising new avenues for pretesting. Meanwhile, we continue to rely on the partnership with our test centres and will keep on looking for ways to add to the service we offer them through, for example, improved feedback and prompt turnaround of results.

Reference

Bachman, L F (2004) *Statistical Analyses for Language Assessment*, Cambridge: Cambridge University Press.

This is an extract from *Research Notes*, published quarterly by University of Cambridge ESOL Examinations.

Previous issues are available on-line from www.CambridgeESOL.org

© UCLES 2005 – this article may not be reproduced without the written permission of the copyright holder.