

The IELTS Question Paper Production Process

The production of IELTS question papers is a lengthy process which includes a number of quality checks. The objective of these checks is to ensure that the material in each test is suitable for the test purpose in terms of topics, focus, level of language, length, style and technical measurement properties.

We apply both qualitative standards for the production of test material involving the judgement of qualified professionals, and quantitative, statistical standards for the selection of suitable test material and the maintenance of consistent levels of test difficulty over time.

The stages in the process of producing question papers are shown in Figure 1 below. The first three stages of *commissioning*, *pre-editing* and *editing* involve gathering and choosing appropriate test content that reflects the aims of the Academic and General Training modules.

Once the best material has been selected, it is then given to representative groups of language learners to check that each question – or *item* - is at an appropriate difficulty level for IELTS; that candidates will be able to understand the questions and that each question can help us to differentiate between more and less able candidates. This stage is known as *pretesting*. Approved material is stored in an *item bank* and can then be introduced to *live* tests – tests that are used as the basis for awarding official IELTS certificates – through a process known as *standards fixing*. Each of these stages is explained in more detail below.

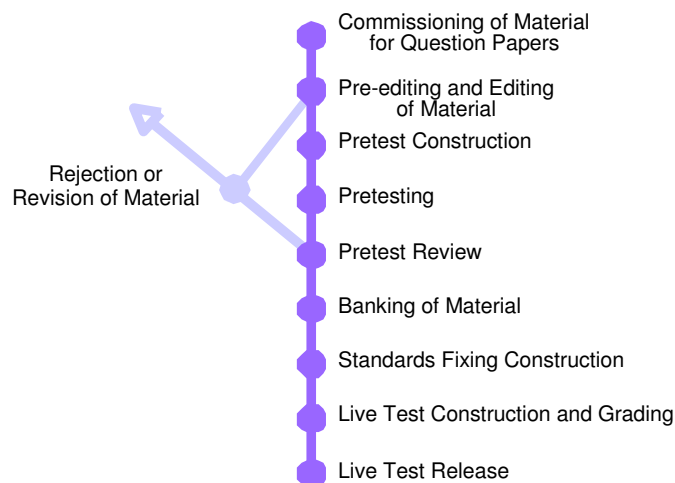


Figure 1. The Question Paper Production Process for IELTS

Commissioning

There are one or two commissions each year for each of our item writing teams. These feed material into the question paper production process. To reflect the international nature of IELTS, test material is written by trained groups of item writers in the United Kingdom, Australia, New Zealand and the USA and is drawn from publications sourced anywhere in the world. Overall test content is the responsibility of both externally commissioned language testing professionals – the *chairs* for each of the Listening, Reading, Writing and Speaking sub-tests – and of Cambridge ESOL staff.

Item writers work from test specifications. These specifications detail the characteristics of the IELTS sub-tests, outline the requirements for commissions and guide writers in how to approach the item writing process including selecting appropriate material; developing suitable items and submitting material for pre-editing and editing.

Pre-editing

Pre-editing is the first stage of the editing process and takes place when commissioned materials are initially submitted in draft form by item writers. A meeting is held involving chairs and Cambridge ESOL staff to review the material.

The purpose of pre-editing is to ensure that test material is appropriate in terms of:

- topic
- topicality
- level of language
- suitability for the task
- length
- focus of text
- style of writing
- focus of task
- level of task.

At this stage, guidance is given to item writers on revising items and altering texts for resubmission. This is seen as an important element in item writer training and advice is also offered on any rejected texts and unsuitable item types.

Editing

Following pre-editing feedback, material is completed and submitted for editing. Editing takes place at meetings involving Cambridge ESOL staff and chairs. Item writers are encouraged to participate in editing meetings dealing with their material. This is seen as another important part of their ongoing training.

At editing, texts and selected items are approved for pretesting or are sent back to a writer for further revision. Revised material is then re-edited at a subsequent meeting.

Pretest construction and Pretesting

IELTS pretests are very similar to the tests that will be used in live administrations. The tasks are in their final form including task rubrics (instructions) and examples. Listening pretests are professionally recorded to ensure that they are of acceptable quality. Listening and Reading pretests are administered to IELTS candidates at selected centres or to prospective candidates on IELTS preparation courses. The pretests are marked at Cambridge ESOL and statistically analysed. Writing and Speaking pretests are administered to representative samples of candidates to assess the appropriateness of this material for use in live tests, and to establish that the tasks are capable of eliciting an adequate sample of language to allow for the assessment of candidates against the scoring criteria.

Pretest Review

The Validation Unit at Cambridge ESOL collates and analyses the pretest material.

Listening and Reading pretests

All candidate responses are analysed to establish the technical measurement characteristics of the material, i.e. to find out how difficult the items are, and how they distinguish between stronger and weaker candidates. Both classical item statistics and latent trait models are used in order to evaluate the effectiveness of the material. 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 is used to locate items on the IELTS common scale of difficulty. In addition, the comments on the material by the staff at pretest centres and the immediate response of the pretest candidates are taken into account.

At a pretest review meeting, the statistics, feedback from candidates and teachers and any additional information are reviewed and informed decisions are made on whether texts and items can be accepted for construction into potential live versions. Material is then stored in an item bank to await test construction.

Writing and Speaking pretests

Separate batches of Writing pretest scripts are marked by IELTS Principal Examiners and Assistant Principal Examiners. At least two reports on the task performance and its suitability for inclusion in live versions are produced. On the basis of these reports, tasks may be banked for live use, amended and sent for further pretesting or rejected.

Feedback on the trialling of the Speaking tasks is reviewed by experienced examiners, who deliver the trialling tasks, and members of the item writing team who are present at the trialling sessions. The subsequent reports are then assessed by the paper chair and Cambridge ESOL staff.

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.

Standards Fixing and Grading

Standards fixing ensures that there is a direct link between the standard of established and new versions before they are released for use at test centres around the world.

Different versions of the test all report results on the same underlying scale, but band scores do not always correspond to the same percentage of items correct on every test form. Before any test task is used to make important decisions, we must first establish how many correct answers on each Listening or Reading test equate to each of the nine IELTS bands. This ensures that band scores on each test indicate the same measure of ability.

Once we are satisfied with the quality of the material, each new test task is introduced as part of a live test administration (with limited numbers of candidates and under tightly controlled conditions). We use information from this exercise to confirm our estimate of how difficult the new task is when compared to the established test material. The task is then ready to be used in combination with other material as part of a fully live test.

Test construction

At regular test construction meetings, Listening and Reading papers are constructed according to established principles. Factors taken into account are:

- the difficulty of complete test versions and the range of difficulty of individual items
- the balance of topic and genre
- the balance of gender and accent in the Listening versions
- the balance of item format (i.e. the relative number of multiple choice and other item-types across versions)
- the range of Listening/Reading skills tested.

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

Data are collected routinely from live administrations and analysed both to confirm the accuracy of the initial grading process and to support additional investigations into quality assurance issues.

Ongoing Research and Development

In addition to this routine of test development and validation, the IELTS partners carry out academic research to support the tests and sponsor external researchers. Details of this research are given in the *IELTS Annual Review*, which can be accessed on the IELTS website: www.ielts.org. Based on this research work, regular improvements are made both to the test itself and to its administration.