

LanguageCert International ESOL

Maintaining Standards

# LanguageCert's approach to maintaining standards

## The International English for Speakers of Other Languages (IESOL) exam suite

The IESOL (Listening, Reading, Writing) and IESOL (Speaking) examinations, along with all LanguageCert assessment products, are high quality tests produced by long-standing experts in English language assessment and the CEFR. Test specifications have been designed to fully reflect the requirements of the CEFR and test materials writers are of the highest international standards and have extensive expertise in, and knowledge and understanding of, the CEFR.

In order to ensure high quality, validation of the levels of the tests is monitored through ongoing independent external research. For example, LanguageCert recently worked with the Centre for Research in English Language Learning and Assessment (CRELLA) of the University of Bedfordshire to carry out an empirical investigation linking the IESOL Communicator examination to the CEFR. International expertise was drawn upon in this highly detailed study that established that the LanguageCert B2 examination (IESOL Communicator) does indeed accurately reflect the contents and ethos of the CEFR.

In addition to such longitudinal and high-level research, LanguageCert employs highly sophisticated item-banking techniques accompanied by detailed statistical analysis of test items. This work ensures that test items are at the correct level of difficulty for the CEFR level tests they are used in and that the difficulty level of the multiple test versions at each level always require the same standard to gain a Pass. Rasch statistical analysis is used alongside more traditional and standard test item analyses in order to ensure that Item Response Theory forms a central part of the item-banking calibration techniques.

Ongoing validation work to ensure that the LanguageCert tests measure appropriately with regard to the CEFR is a vital component of the work of LanguageCert and it is this effort that ensures that the examinations remain fully fit-for-purpose and aligned with the CEFR system of levels into the future. It is our belief that no examination system is ever perfect and that technological advances and developments in language acquisition theory need to be taken into account in changes to examinations to ensure best fit with the CEFR and modern thinking on assessment.

## Validity and the LanguageCert IESOL exams

The LanguageCert International ESOL tests are designed to ensure fitness for purpose and to deliver assessments which take into account contemporary views on validity. Validity is generally defined as the extent to which a test measures the intended purpose, in this case communicative English language proficiency. It is concerned with appropriateness and meaningfulness of results. The individual qualities of validity and reliability need to be considered together in order to ensure fairness to candidates and to generate trusted result outcomes that will replicate real-world performance of candidates. The LanguageCert tests are designed in this way.

The construct of a test is the theory that the test is based on. For the LanguageCert tests this is the theory of communicative language ability. The approach to language testing in these tests is based on models of communicative language ability (CLA) such as Bachman (1990) and the foreign language specifications provided by the Council of Europe in such documents as Waystage (1990) and Threshold (1990). LanguageCert tests see language proficiency in terms of language users' overall communicative ability subdivided into skills and sub-skills. The tests measure each of the four language skills – Listening, Reading, Writing & Speaking – separately, as they can be recognized and assessed separately. Each test provides an individual profile per skill, and these can then be added together to give an overall picture of a candidate's language proficiency.

Construct validation activities are carried out by LanguageCert beginning with test and task design. Experts analyse tasks and content in an ongoing manner to ensure they are fair, have interactional authenticity and sample the appropriate language skills for the level and skill. Cognitive processes on the part of candidates taking the test also ensure authenticity, as well as the sampling of real-world skills. This ensures the cognitive validity of the tests and that the Target Language Use (TLU) domain (Bachman and Palmer, 1990) is considered. Validation of this area is carried out in terms of analysis of whether the tasks and items are representative of the TLU. The LanguageCert tests employ a wide variety of task types to ensure a wide range of language is elicited and skills tapped into, and so provide the maximum evidence of the underlying skills and abilities tested. Materials writers are highly experienced and expert in this area. Context-related validity concerns the conditions under which tests are taken and includes not only the nature of tasks, rubrics and topics, but also actual exam-taking conditions. Regular and robust investigation of these areas ensures the validity of the tests is maintained.

The test development process of LanguageCert English language exams has been established to ensure Validity is achieved. Bachman (1990) states that language tests should support inference to some domain of 'Target Language Use' (TLU). That is, in order to judge the validity of test results, we must first state what we expect a test-taker to be able to do in real-world language use. The Common European Framework of Reference (CEFR) has been utilized to help determine the test construct of the LanguageCert exams for this purpose. Its illustrative descriptors across a range of language domains and contexts have been used as a starting point and extensively inform the test development processes employed.

The task types used in the LanguageCert examinations have been selected to ensure they have interactional authenticity and can be related to real-world performance. They directly sample the cognitive skills, strategies and language knowledge that support inference about the potential ability of a candidate in real-world interactional situations.

For example, in the receptive skills, Reading tests are carried out using real-world notices and signs that students need to show understanding of. Reading also focuses on aspects of comprehension of texts at the appropriate level using a variety of task types, such as multiple-choice selection, matching and open response to prompts. The degree of challenge and sub-skills required in the Reading tests start with straight tests of understanding at A1 level, through to text interpretation and understanding writer intention at the B levels, through to quite complex interpretation strategies required at the C levels to infer the meaning of unknown lexis from context and the ability to sift through lengthy reading inputs to identify key information and to show awareness of text cohesion.

Similar skill sets are tested in Listening to ensure interactional authenticity. At the lower levels of proficiency, candidates need to be able to pick out key information from short recordings and conversations of the type they would participate in in the real-world. At the higher proficiency levels, candidates need to follow lengthy exchanges to pick out key points and understand speaker intention and nuance. All tasks focus on interactions that candidates will need to engage with in the real world.

With the productive skills of Writing and Speaking the focus is more on use, and validity evidence relates to language actually being used for a variety of communicative purposes. Tasks are explicitly designed to ensure the candidate engages in real-life situations, whether this is writing to a friend or making a complaint, through to direct interaction with an examiner in the Speaking test. In both the receptive and productive skills, the emphasis is always on interactional authenticity. The tasks are designed to ensure candidates engage in the same type of activity and engage similar cognitive processes and language as would be required in a real-world language activity.

Validity thus links performance on the tasks in LanguageCert International ESOL tests to an inference about the test taker's ability in a world beyond the test. The tests are designed to elicit a sample of performance which is interpretable and generalizable to the real world. In order to ensure the test results are generalizable CEFR Can-Do statements have been used as the basis for what test-takers need to be able to achieve at each level.

## Achieving Reliability

Reliability relates to consistency in test results. This is achieved in the LanguageCert International ESOL tests by ensuring test forms are comparable in terms of content and difficulty, and through robust item-banking techniques, involving the pretesting and trialling of test materials and the placement of all items on the LanguageCert Item Difficulty (LID) scale.

Reliability is crucial for all test stakeholders who need to be sure that different administrations of the test deliver very similar results. This is essential for fairness to test-takers and to ensure that receiving institutions such as universities and employers can be guaranteed that the same ability level is required to pass the same examination at different administrations. The start of the process of ensuring reliability of results is to ensure standardisation of test-taking experience. This begins with test specifications that ensure tests can be replicated over years of administrations, through standardised test-taking conditions and finally through the difficulty of the test materials and the way tests are graded.

Specifications and robust standardised item-production techniques permit a constant supply of new test items into the item bank. Harmonised procedures for test day administration are provided to test centres. Item-banking techniques, using statistical analysis provided from pretests and live tests ensure empirically that tests are always of very similar difficulty.

## Impact and Fairness

Assessment has important effects and consequences for a range of stakeholders within the area of education, and also within society more widely. Test-takers in particular are affected because the results of English language tests are used to make important decisions which can affect their lives. All the LanguageCert International ESOL examinations aim to have a positive impact on test-takers and on English language education. This approach is coherent across the whole test development process. All stages from determining test specifications through to delivering results put impact on learners at the centre of the work of LanguageCert. LanguageCert also gives high importance to the positive washback of the examinations on the classroom, where the learner should be engaged in meaningful, communicative tasks in order to succeed in real-life communication and in the LanguageCert tests.

Assessing candidates' performances can be a complex and challenging task. LanguageCert is committed to approaching this challenge in a robust and responsible manner, implementing assessment methodologies which are up-to-date and appropriately and accurately reflect candidates' language proficiency and do not present assessors and/or learners with any unnecessary practical impediments.

LanguageCert takes very seriously the requirement to ensure assessment outcomes that are comparable in standards between assessments within a qualification, between similar qualifications and with other awarding organisations. This is critical in providing fit-for-purpose, fair tests for candidates and other stakeholders, and in ensuring that results and standards are fully comparable over time.

LanguageCert puts fairness to candidates at the very heart of all its work. All Assessment staff strive to ensure that tests are valid, reliable and have a positive impact on learners. An important part of ensuring fairness to test-takers is to minimise any bias in the test materials. The process of eliminating bias begins with the formation of the test specifications. These are written with direct reference to the nature of the intended or anticipated candidature to ensure the tests are fully fit-for-purpose. LanguageCert makes sure writers understand who the target test users are, and that they consider aspects such as the level of cognitive processing of typical candidates, and also the cultural contexts they will be used to.

The requirement to produce materials which will not favour or discriminate against certain candidates is central in LanguageCert's item development process. This includes ensuring test materials are as free from specific regional or national cultures as possible, and that topics are universal. A list of taboo topics is provided to item-writers to aid in this. These taboo topics include areas which may cause distress or distraction to candidates or relate to unfortunate experiences they have suffered through to specific aspects of local cultures which may be alien to the local culture of the candidate or beyond their life experience. The LanguageCert team are also extremely careful not to introduce test material which may test general knowledge or specific technical knowledge rather than language ability.

Through these very thorough and robust procedures, LanguageCert takes the issue of minimising bias very seriously. Minimising bias is about ensuring that an assessment does

not produce unreasonably adverse outcomes for learners who share a common attribute. The minimisation of bias is related to fairness to all learners and is also closely related to statutory equality regulations. All LanguageCert exams are designed to be fair.

LanguageCert is a business name of PeopleCert  
Qualifications Ltd, UK company number 09620926.

Copyright © 2019 LanguageCert

All rights reserved. No part of this publication may be reproduced or transmitted in any form and by any means (electronic, photocopying, recording or otherwise) except as permitted in writing by LanguageCert. Enquiries for permission to reproduce, transmit or use for any purpose this material should be directed to LanguageCert.

#### DISCLAIMER

This publication is designed to provide helpful information to the reader. Although care has been taken by LanguageCert in the preparation of this publication, no representation or warranty (express or implied) is given by LanguageCert with respect as to the completeness, accuracy, reliability, suitability or availability of the information contained within it and neither shall LanguageCert be responsible or liable for any loss or damage whatsoever (including but not limited to, special, indirect, consequential) arising or resulting from information, instructions or advice contained within this publication.