



The global forum for
EAP professionals

*BALEAP Guidelines on
English Language Tests
for University Entry*

Contents:

Introduction	2
Who these guidelines are for	2
Selecting assessments and setting levels for course entry	3-4
Test Digests	9+
IELTS	
TOEFL	
PTE Academic	
MELAB	
TEEP	
CPE, CAE (Cambridge Main Suite)	
BEC	
School and Vocational exams	

BALEAP is a professional organisation whose members are providers of English for Academic Purposes (EAP).

Its aims are to :

- enhance the quality of English language provision for international students in institutions of higher education and to support professional development of the staff
- provide an accreditation scheme for EAP courses
- promote and disseminate EAP associated research through biennial conferences and one-day professional Issues Meetings (PIMs) and also through publication of research and conference presentations.

This guide was compiled by the BALEAP Testing Working Party: Bruce Howell (University of Reading), Philip Nathan (Durham University), Diane Schmitt (Nottingham Trent University), Chris Sinclair (University of Southampton), Jenifer Spencer (EAP writer and editor, formerly Heriot-Watt University), and John Wrigglesworth (University of Portsmouth). They bring together close to 100 years of EAP experience.

Introduction

The purpose of the *BALEAP Guidelines on English Language Tests for University Entry* is to give stakeholders a richer description of English language tests¹ in a format that facilitates comparison between them. Our aim is to assist staff responsible for reviewing and selecting the tests used by their institutions and setting the related scores/grades. It also aims to assist in the day to day interpretations of scores on the qualifications concerned.

Who are these guidelines for?

Relevant stakeholders including all those responsible for setting and using English language entry requirements in Higher Education institutions, including:

- HE institution admissions officers and their supervisors
- International office and marketing staff
- EAP administrators and course leaders
- EAP teachers
- Sponsors, agents and students themselves

It should be noted that these pages are **guidelines** and not definitive declarations. Institutions need to make their own decisions about admissions policies and it is strongly recommended that admissions supervisors engage with all relevant staff to discuss the issues raised here, as well as to disseminate policy.

Why is there a need for a guide?

A fundamental aspect of language testing or assessment is the use to which test scores are put. Therefore, the end user of a test score is as important as a test developer or administrator in ensuring that a test is valid. It is, thus, important that all test users develop a degree of understanding of the relationship between test purpose, format and the meaning of test scores in order to set and apply realistic and fair standards.

For the most part this comes down to knowing what tests can and cannot do and considering how language test scores should be used in conjunction with other evidence to enhance the dependability of admissions decisions. When considering the suitability and grading requirements of any particular test, the task of stakeholders can be facilitated by routinely applying a checklist of questions, such as the one on page 2 below.

¹ the terms 'test' / 'examination'/'exam' are used interchangeably in this guide

What should we consider when selecting assessments for course entry?

Selecting appropriate qualifications for entry and the relevant minimum grades is highly situation-dependent. Differences between institutions, disciplines and the pedagogies of courses mean that each will make different demands on the students' language ability, study skills and content knowledge. Those responsible for selecting tests thus need to be aware of the particular language needs of their own students rather than thinking in terms of a prototypical university student.

Language tests also differ in scope, structure and scoring, so that the range of language that is assessed will differ from test to test. This makes direct test and score comparisons difficult. Adopting an entry requirement from a different type of course or from a different institution is therefore not advisable. Because we rarely know the procedures other institutions or programmes have followed in selecting tests and related scores or the details of their courses, there is a danger of circularity where the only justification for English entry requirements becomes 'that's what everybody accepts'. We strongly discourage this circular approach.

For these reasons, we recommend that advice on tests and admissions criteria should be taken and applied **locally** and reviewed annually. We recommend that admissions staff, academic departments and EAP practitioners work together to monitor and track the results of the decisions that they make in order to reflect on and keep a record of the effects of their policies.

What are the key questions?

The following is a list of important questions which can be used as a basis for developing your institution's own checklist for adopting any test as suitable evidence of language proficiency:

- What elements of language does the test evaluate? (i.e. the UKBA requires evidence of proficiency in all four skills)
- Does it actually test what it says it tests? What is the test content and what are the assessment tasks?
- How long is the test? (Most tests for this purpose take approximately three hours. A shorter test will be limited in the amount of information it can provide.)
- Does the score report or certificate tell us what we need to know? (Does it provide a breakdown of scores by skill or simply an overall mark?)
- How long is a score valid?
- At what range of scores does the test discriminate most reliably, in terms of giving an accurate indication of a student's proficiency?

- What is the evidence that students accepted at a certain grade on this test have sufficient language proficiency to perform satisfactorily on their course of study²?
- Are the processes for applying to take the test, sitting the test and presenting the score report secure?
- What measures are taken to ensure the consistency and reliability of the setting and marking of individual versions of the test?
- How is the score reported to my institution? Is there a reliable and convenient way of checking the legitimacy of scores?
- Does the relevant pool of applicants have sufficient access to test centres?
- What types of language support are available for international students in my institution?

How do we go about setting a minimum entry score?

This process is called standard setting or setting cut scores. It may be undertaken at an institutional level or on a programme by programme basis. Most importantly, it should be viewed as a local process. Test providers do NOT set entry requirements, although many (e.g. TOEFL, IELTS, PTE) do provide guidance on how institutions might go about setting cut scores. For specific guidelines for an individual test contact the test provider directly or go to their website. A cut score will need to be set for each test that you choose to accept for admission.

The standard setting process normally involves putting together a panel of stakeholders ideally made up of:

- admissions staff who will process applications
- academic staff from programmes that receive international students,
- marketing and international staff who can supply information on recruitment targets and relevant pools of applicants
- English for Academic Purposes staff who can offer guidance on language learning and language support opportunities in the institution

To set standards, the panel must consider what students need to be able to do with language and the minimum level of the language required to carry out these activities. Although all university students will need to speak, write, listen and read to some degree in all programmes, the emphasis may differ considerably from one programme to another or across levels of study. This first step should be done independent of any specific test.

² Note that predictive validity really only works for the first semester or so. After that, too many other variables come into play to determine that success or failure is solely due to language proficiency.

Next, the panel should read or listen to sample student performances on various test tasks. They should select which student performances meet the minimum standards set in step one. They then compare these performances to actual test scores received and use this information to set a cut score. This second step is normally an iterative process that involves plenty of discussion between panellists until agreement is reached. In their deliberations over where to set cut scores, panellists should also take account of the availability of in-session language and skills support for students across the university or on particular degree courses. However, the availability of in-session support should not be used as justification for setting low entry scores.

After completing steps one and two, the panel may recommend that their institution set one cut score for all applicants or that it set a variety of cut scores according to the profile of skills required by particular programmes. For example, a university might set a high score for speaking for students studying in programmes that require a lot of group work or a work placement and a lower score for speaking for students undertaking research degrees.

Including panellists from across the university, ensures that any discussion and decisions take account of factors relating to language learning, academic requirements and information about the international student market. Setting entry requirements that are too lenient can lead to high drop-out levels and student dissatisfaction, not only from the students who find they have inadequate language skills for the course they are undertaking, but also from peers who feel that this hinders course delivery and their own learning experience. On the other hand, setting the bar too high might result in both students and the institution losing out if an able student is unnecessarily barred from entry.

How should minimum entry scores be applied in practice?

When using cut scores, it is important to keep in mind that language test scores have been characterised as exhibiting “inevitable uncertainty” because a range of factors (e.g. tiredness, stress) prevent any test from being truly precise in its measurement of any person’s language proficiency. Thus, a test score can only give us an approximation of any test taker’s “true” proficiency level. In recognition of this, test providers normally calculate and report a statistic called the Standard Error of Measurement (SEM). The SEM represents the level of confidence that a score user should have in an individual test score. The smaller the SEM in relation to the length of the scale the more reliable the test. A test with a large SEM means that score users can have less confidence in how to interpret any individual score.

The following example illustrates how an admissions office might use the SEM in the process of making an admissions decision. The SEM for an overall score on the iBT TOEFL is +/- 5 points. If a university has set a minimum entry score of 87 and receives an otherwise excellent application from a student with an iBT TOEFL score of 84, a strict interpretation of the cut score would require that the student be rejected. A score interpretation that takes account of the SEM for iBT TOEFL (in this

case treating the minimum score as a range between 82-92) alongside all of the other evidence supporting this student's application may lead to a recommendation that the student be accepted. In line with the advice from most testing organisations, we advise that score users always interpret test scores alongside other admissions evidence and with the SEM in mind.

What is the CEFR?

The Common European Framework of Reference for Languages – Learning, Teaching and Assessment (CEFR) is a document which sets out to describe what learners have to do to use language at various proficiency levels and in various contexts. One aim of the CEFR is to provide a profile of language users' abilities across different types of language use, e.g. reception, interaction and production. As each of these types of language use can be assessed in a variety of ways by different exams, a further aim of the CEFR is to facilitate mutual recognition of language qualifications across Europe in terms of language learning objectives and language learner achievement.

Thus, when an examination or test is linked to the CEFR, the test provider uses the common language of the CEFR performance descriptors to provide an indication of what test takers with particular scores can do. The CEFR descriptors do not, however, provide information about any given test's original purpose or the types of tasks used by any test to assess learner language.

This common approach to describing language tests in no way implies that all tests that are linked to the CEFR serve the same purpose or provide the same amount or quality of information about a language learner's ability to perform in a particular language use domain. Stakeholders must refer to documentation provided by testing organisations about the purpose, content, validity and reliability of their tests. Only when CEFR descriptors are used alongside information about specific exams and language use contexts do the descriptors become truly meaningful.

How do UKBA CEFR requirements affect the entry levels set by my university?

The UK Border Agency has set English language entry requirements that they feel are necessary to achieve the goals of immigration policy. Your university needs to set English language entry requirements that match its own goals for international recruitment and international student success. Standards for success at university should be set in relation to the requirements of your own institutional standards and requirements.

The UKBA use of the CEFR levels enables them to refer to language proficiency levels independent of any single test. As long as entry requirements are equal to or higher than those set by the UKBA, universities remain free to set their own English language entry standards. In most cases, the level of English required for success at university will exceed the minimum requirements set by the UKBA for immigration purposes.

Please refer to the most up to date UKBA Tier 4 of the PBS Policy Guidance for information on the CEFR levels required by the UKBA for immigration purposes. <http://www.ukba.homeoffice.gov.uk/sitecontent/documents/employersandsponsors/pbsguidance/>

How do UKBA requirements affect the exams or types of assessments of English language that we can accept?

The UKBA has prepared a list of approved English language tests called Secure English Language Tests (SELTs). This list was compiled for use with all of the different immigration tiers. Therefore, not all of the tests are suitable for all immigration purposes. For example, some of the tests are not able to provide reliable scores for immigrants needing only low levels of English proficiency, while others do not include content which enables assessment of an applicant's readiness to study at university through the medium of English. Each test on the list should be considered in relation to its purpose and evidence which demonstrates its validity for that purpose.

At present, universities are not limited to using SELTs as evidence of applicants' English language proficiency. Universities can use any assessment of their choice provided applicants are assessed on all four skills – reading, writing, listening and speaking and have achieved CEFR B2 level proficiency. It is the responsibility of the university to ensure that the assessment used provides a valid assessment of prospective students' language proficiency and that there is demonstrable evidence of how the test has been linked to the CEFR. Universities who have their own in-house assessment are encouraged to use the same procedures as those used in these guidelines to ensure that their own assessments are fit for purpose.

If students require a pre-sessional programme, how do we determine how long it should be?

Unfortunately, this is an area where practice is led more by market forces and anecdotal evidence rather than research. It is extremely difficult to state categorically that X number of teaching hours will lead to a specific score gain on a particular test because of individual differences between learners and differences in learning conditions. Thus, it is important to bear in mind that publicised estimates of the number of hours required to move from one level of proficiency to another are not supported by research evidence.

Those responsible for designing and setting entry requirements for pre-sessional language courses, should take account of the wide variability between learners and should not foster expectations of rapid progress in short periods of time no matter how intensive the provision may be.

An essential tool in making sure that the language levels set are realistic and effective is to have formal tracking procedures. This involves following a number of students (the more the better, and from a variety of backgrounds) from arrival

through to graduation. Data collection should include English language test scores, marks on degree assignment and rate of progress through the degree. This quantitative data should be supplemented with interviews with tutors and the students themselves. An appraisal of the student's ability to cope or succeed on their course can then be related back to their original language qualifications and generalisations for future students can be made. For further advice on conducting tracking studies please go to the BALEAP website.

Test Digests

The next section provides digests for seven tests frequently presented as evidence of English proficiency for university entry. Some of these tests have been specifically developed to assess the language proficiency of applicants applying for university entry e.g. IELTS, TOEFL, PTE Academic, MELAB and TEEP. Others were originally developed for other purposes, but are now marketed by the testing organisation as suitable for assessing readiness to study at university. There is also a section with comments on the use of some professional or vocational exams and school examinations, both domestic UK examinations and overseas school examinations which may be presented as evidence of language proficiency.

Digest format and contents

To help readers to make more informed decisions about the uses to which each test can be put, we have summarized key information in tables. These cover the test length, structure and content including the language, and skills tested. Each digest is followed by a critique based around the following sets of questions:

Validity: Does the test content (in terms of topics, language and skills) assess what the exam claims to test? Is the test content relevant to the academic context in which students will operate? For example, does the writing task require examinees to synthesize information from texts they have read or listened to, thus reflecting an academic assignment, or is the writing based only on examinee's personal opinions?

Reliability: This relates to the procedures used to ensure that the construction, delivery, and marking of an exam produces consistent results, and includes a review of the training and monitoring of examiners and markers. Test providers should carry out on-going empirical research into candidate performance across different administrations and with different test taker populations and produce publicly available reports on their findings.

Test security: Fraud among test takers is well-documented whether it be cases of the impersonation of examinees by substitutes or the presentation of counterfeit certificates. We have therefore included information on test providers' security measures at the point of test administration (e.g. the use of photo ID) and also in providing results in a way that can be independently verified, directly from the testing organisation.

Overall Evaluation: These guidelines aim to make clear that English proficiency tests cannot be directly compared with each other. When deciding on entry requirements those involved in admissions decisions need to weigh carefully the strengths and weaknesses of each test and its appropriacy as an entry requirement for a particular course of study. Obviously language requirements vary, depending on the subject matter. However, it is important not to underestimate the demands on language even for more mathematical and practical subjects. The evaluative comments on each test are based on the principles developed by the BALEAP working party on testing based on our collective experience and research.

Guidelines and advice on evaluating additional tests which may be offered as proof of English proficiency

We hope that the format of the test digests will provide a helpful template when evaluating other tests. We advise that admissions departments should contact test providers to request information on similar lines to that provided in these digests. It should be clear that any test whose providers cannot supply verifiable evidence (e.g. copies of the relevant reports) of reliability and security should not be considered appropriate for high stakes direct entry purposes.

When assessing the validity of a test for academic purposes, thought needs to be given to the suitability of the exam content, context and even length. Tests which rely on sentence level exercises and multiple choice items are unlikely to adequately assess a student's preparedness for writing extended academic assignments. A one-hour test is unlikely to provide satisfactory evidence of the student's range of skills. However, different considerations may be appropriate when judging such exams for indirect entry to 'pathway' courses such as pre-sessionals.

The IELTS test is developed and administered by The British Council, IDP Australia and Cambridge ESOL, and is the most frequently submitted English language test for entry to higher education in the UK and Australia". There are two IELTS tests, the Academic and the General Training test, only the academic test is intended to be suitable as a measure of English proficiency for entry to Further and Higher Education academic programmes.

Academic Test Characteristics

<p>Skills tested:</p> <p>Exam length: Approximately 3 hours Mode: Paper-based and audio listening (R/L/W2hrs 45 minutes-speaking administered separately)</p>
<p>Reading</p> <p>Time: 60 minutes. Three passages: 2000 – 2700 words in total. Narrative, descriptive or discursive texts are taken from books, journals, magazines and newspapers. Around 40 test items carrying one mark each, including multiple choice, short-answer questions, identifying information, writer's views/claims, matching information, headings, features, or sentence endings; completion of sentences, notes/diagrams and table summary completion.</p>
<p>Listening:</p> <p>Time: 30 minutes plus 10 minutes transfer time. Mode: Audio</p> <p>Four sections: The first two texts, a conversation and a monologue, involve everyday social contexts. The other two texts involve educational or training contexts. The final text is a monologue on an academic subject. The recordings are heard once only. There are around 40 test items, with the answers in order of listening. Tasks comprise a similar mix to those for reading.</p>
<p>Writing:</p> <p>Time:60 minutes</p> <p>Task 1: Time: 20 minutes</p> <p>Candidates are asked to describe visual information (Graph/diagram)at least 150 words. This may involve some comparison or identification of trends. However candidates are not expected to speculate on possible causes or significance of the data.</p> <p>Task 2: Time: 40 minutes.</p> <p>Essay (at least 250 words) In response to a prompt in the form of an issue or problem or statement of a stance. Student answers are non-specialist and based on personal opinion and experience. This task assesses the candidate's ability to present, organise and support their ideas in the form of a supported argument and assesses language accuracy and range.</p>
<p>Speaking:</p> <p>Time:11 – 14 minutes</p> <p>Format: Three parts; Oral interview between candidate and examiner: Long turn on a topic from a prompt card - 1 minute preparation, 1-2 minutes extended speaking. Candidate and examiner conduct a discussion of the same topic, in a more abstract and general way. (4-5 minutes. Content is not academic. Taken on the same day or up to seven days before or after the main test.</p>
<p>Integration of skills: Candidates are not required to integrate the different skills in any tasks.</p>

<p>Scores: A global score between 1 and 9 is awarded, with scores also recorded as a profile, on the same scale, for each module of the test. Scores are reported in full and half bands, e.g. IELTS 6.0 or 6.5. Scores are available within two weeks of the candidate taking the test and are seen as valid for no longer than two years. There are no restrictions on the number of times a candidate can retake the test.</p>
<p>Scores & scoring procedures Reliability scores for reading and listening (the objectively marked modules) are produced annually. Reliability for the writing and speaking are ensured through explicit criteria, benchmarking and examiner evaluation training, Research on the IELTS is available under the Research tab on the IELTS website.</p>
<p>Test security Procedures to ensure candidate identity including photographing candidates. Online verification of results available (TRF).</p>
<p>Availability & accessibility 6,000 centres worldwide. Offered four times per month in 250 countries. Fees are collected in local currency and are set locally.</p>
<p>Test preparation and test practice resources: Practice papers for students and sample lessons for teachers can be downloaded from the website. Also a catalogue of official published materials including Handbooks for teachers, Official IELTS Practice Materials, Past Paper Packs and research information. There are also a wide variety of commercially produced IELTS preparation books.</p>
<p>Contact details: www.ielts.org</p>

Evaluation & Comment

Although IELTS is the most widely recognized and used and test for direct entry to UK Higher Education, care must be taken in using the test appropriately. The test is seen as most discriminating between Bands 5 and 7; thus, it is best used for discriminating between students who are ready for pre-sessional entry courses and those who may be ready for direct entry. The relationship between IELTS and the CEFR is not entirely clear. The test providers point out that “As IELTS preceded the CEFR, IELTS band scores have never aligned exactly with the CEFR transition points”. On these grounds, they currently suggest that institutions should set a requirement of Band 7 rather than Band 6.5, if they require a high degree of confidence that the applicant is at C1. Further information about IELTS and the CEFR is available at http://www.ielts.org/researchers/common_european_framework.aspx

Although they cite Hawkey and Barker³ and internal research notes, they point out any correspondences suggested are also based on experience and that research is incomplete. The global test score should be read along with the score for each skill with an eye for differences in the component scores (a jagged profile). A range of reports are available on validity, reliability and impact these are available at <http://www.ielts.org/researchers/research.aspx>.

Some concerns which have been raised about the test format may be the authenticity of some tasks, in terms of an academic context. For example, Writing

³Hawkey, R. and Barker, F. (2004) Developing a common scale for the assessment of writing, *Assessing Writing* 9 (2), 122-159.

Task 1, in which students are only asked to describe a visual representation, but not asked to give any suggestions or explanations about the data. Task 2 is also the type of essay based purely on opinion that would be less common in academic contexts. The reading texts are realistic in length, but the lack of integration of skills means that the tasks do not generally require the type of authentic responses which might be needed in an academic context.

TOEFL - Test of English as a Foreign Language **Test Digest**

TOEFL is developed and administered by ETS, Educational Testing Service, a US-based non-profit organisation. TOEFL is the most frequently submitted English language test for entry to higher education in North America, and is also accepted at many academic institutions worldwide.

Academic Test Characteristics

Skills tested	
Exam length: approximately 4 hours	Mode: Entire exam is computer based.
Reading:	
Time: 60- 90 minutes	
3-5 passages 700 words long each	
The passages are excerpts from introductory sections of college-level textbooks	
12-14 questions for each passage, including prose summary completion, table completion and multiple choice.	
Listening:	
Time: 60-90 minutes	
4- 6 excerpts from lectures, some with classroom discussion, 3-5 minutes long with six questions	
2- 3 conversations from an informal academic context - each 3 minutes long. Five questions each. Answers are in the form of chart completion and multiple choice.	
Writing:	
Time: 50 minutes.	
Task 1: Time: 20 minutes; integrated task involving Reading/Listening/Writing.	
There is a short academic listening and a short reading passage on the same topic. The students have to describe how the two texts relate - one usually involves some sort of critique of the other. Students can access the reading passage during writing.	
Task 2: Time: 30 minutes; essay response to a question which invites comment based on personal experience or opinion.	
Speaking:	
Time: 20 minutes	
Two independent questions about familiar topics, where responses are based on opinion or personal experience	
4 integrated questions, where candidates read a short passage, listen to a short related text and then integrate the information in their own words. The content is drawn from academic and campus-based material.	
This set consists of two Reading/ Listening/ Speaking questions and two Listening/Speaking questions with responses based on what was read and heard. Response times allowed are	

<p>from 45 to 60 seconds for each response. Students listen to conversations or lectures via headset, see text and context-setting visuals on their computer screen and speak into a microphone. There is brief preparation time, depending on the type of question, and then they speak for approximately one minute. Responses are encrypted and then sent electronically to ETS for scoring.</p>
<p>Integration of skills: Some texts or lecture sources are used as a basis for one of the academic writing and 4 of the academic speaking tasks. The listening and reading sections assess these skills with “stand alone” tasks. Brief additional reading and listening texts on a common topic and having a critical relationship with each other provide the input on which one of the two writing questions is based.</p>
<p>Scores and scoring procedures: Each section is scored on a scale of 0-30, giving a total score scale of 0-120. Section scores and total scores are reported in one-point increments to allow for finer distinctions of ability, according to ETS. Reading and listening are scored directly on a 0-30 scale, but Writing and Speaking are scored as bands 1-6 and then converted to a 0-30 scale. ETS recommends that scores are valid for two years. Students are given a test report with their total score out of 120 and a score out of 30 in each paper. There is also a guidance sheet, interpreting the meaning of their scores in terms of the sub-skills that have been tested and advice for improvement.</p>
<p>Test security: Candidates are photographed and photographs appear on score reports. Students receive a copy of their score report. ETS also sends official score reports with a photograph directly to up to 4 institutions designated by the student, on copy-evident paper or as encrypted electronic files. ETS advises that institutions should never accept score reports provided directly by students as final proof of an applicant’s TOEFL score. On payment, students may designate additional institutions to access score reports, via a verification line, for up to two years after the test is taken, via a password protected account.</p>
<p>Availability and accessibility: The TOEFL Internet-based test (iBT) has been administered worldwide since September 2006. The TOEFL iBT test is offered 30 – 40 times a year at over 4,500 authorized test centres throughout the world. It is available in England and Wales and the Republic of Ireland but not in Northern Ireland or Scotland. The Paper-based (PBT) test is offered six times a year in areas where internet-based testing is not available. The Computer-based (CBT) ended September 2006 and is longer valid.</p>
<p>Contact details: http://www.ets.org/toefl has sections targeted at test takers, institutional users and English teachers. Designated contact email addresses are provided to contact service teams for test-takers, institutions and language teaching providers, respectively.</p>
<p>Test preparation and test practice resources: TOEFL iBT sample questions are available for practice on the TOEFL webpages. ETS/ McGraw Hill publish The Official Guide to the TOEFL Text and a number of major publishers also have test preparation materials available for sale.</p>

Evaluation and Comment

TOEFL is a test with good standards of reliability and security. The marking and test-setting procedures and personnel are vetted and monitored continuously. The test content aims at an academic context through academic related content and the inclusion of integrated skills testing. Extensive research reports and monographs commissioned by ETS on aspects of validity, reliability, impact, institutional and test taker needs can be accessed directly through the Research Link on the TOEFL

website. <http://www.ets.org/toefl/research>. Test and data score data summaries are also published on the website.

The very short response times in the speaking test and the lack of opportunity for extended monologue or interaction make this part of the test a questionable indicator of performance. The rubrics used to assess these responses seem to draw rather wide inferences from such restricted speaking opportunities. However, some teachers would argue that these do assess fluency and that when preparing students for TOEFL the short response times encourage teachers to focus on getting students to respond quickly rather than encouraging the lengthy wait times allowed in EFL settings. Quick responses are required if students hope to participate in seminar discussions or group work. The US cultural contexts of the speaking, listening and writing may cause some difficulties for teachers and students unfamiliar with these contexts.

The use of single point scales in the scoring of all the tests and in the total score may imply more accuracy of discrimination in a test taker's performance than is really possible. Recent research sponsored by ETS⁴ to establish correspondence between TOEFL scores and CEFR levels suggests that the test is likely to discriminate users in the ranges B1 to C1, but would lack discrimination above or below this level. The researchers point out that the TOEFL test was not designed to test CEFR levels, but to assess language use in an academic context. These findings emphasise that care must be taken in using the test appropriately as it is best used for discriminating between students who are ready for pre-sessional entry courses and those who may be ready for direct entry.

PTE Academic-Pearson Test of English- Academic

Test Digest

Pearson Test of English Academic (PTE Academic) is a new computer-based international English language test developed and administered by Pearson Education. The test aims to measure the test takers' academic English language competency in Listening, Reading, Speaking and Writing.

Skills tested

Exam length: Approximately 3 hours

Mode: On-line

(optional 10 minute break between part 2 reading and part 3 writing).

Speaking and Writing:

Total time for both skills: 77-93 minutes

Speaking consists of free speaking in the form of a personal introduction, (not scored but sent to institutions with score report). Scored tasks are: describing an image, such as a map or diagram; integrated speaking tasks: reading aloud a short passage

⁴Tannenbaum, R.J. and Wylie, E.C., Linking English-Language Test Scores On to the Common European Framework of Reference: An Application of Standard-Setting Methodology
TOEFL iBT Research Report TOEFLiBT-06June 2008. ETS, Princeton, NJ: accessed 12.Oct 2010 at <http://www.ets.org/Media/Research/pdf/RR-08-34.pdf>

<p>(up to 60 words, repeating a heard sentence, re-telling a lecture (of about 90 seconds) and answering a short question with a single word or a few words.</p>
<p>Writing: Task 1: Summary: reading and summarising in one sentence an academic style of text of up to 300 words. Task 2: Time: 20 minutes: Essay: Prompt 2-3 written sentences. Write a 200–300 word essay on a given topic.</p>
<p>Reading: Time: 32-41 minutes 5 texts in an academic style, from between 80- 300 words. The tasks consist of multiple choice questions on content and tone of two texts, reordering paragraphs, filling in gaps.</p>
<p>Listening: Time:45-47 minutes There are a variety of audio prompts, in academic contexts and/or styles lasting from 3-5 seconds for the dictation to 90 seconds for the mini lectures. Each is heard only once. Tasks include: writing a 50-70 word summary after listening to a recording (10 minutes), multiple-choice question on the content or tone of the recording by selecting one or more responses, select the missing word from a list of options, selecting the paragraph that best summarizes the recording, highlighting incorrect words in the transcript of a recording, typing a sentence that has been heard.</p>
<p>Integration of skills: The integration of skills is used widely in the test and although the sections are flagged according to the predominant skills focus, individual items are flagged according to the skills involved: e.g. listening and writing, listening and speaking.</p>
<p>Scores and scoring procedures: The score report provides three types of scores: an Overall Score, (range 10-90 points); scores for Communicative Skills (i.e. Listening, Reading, Speaking and Writing), which are based on all items that assess these skills, thus making use of information from the items requiring integration of skills. The range for each skill is 0-90 points. Scores of 0-90 points are also awarded for Enabling skills (i.e. Grammar, Oral Fluency, Pronunciation, Spelling, Vocabulary and Written Discourse). PTE Academic scores are delivered online to test takers, within five business days, via personal login to their on-line account and to registered institutions via their secure login. Test takers can make their scores available to an unlimited number of institutions of their choice. Scores are displayed both numerically and graphically.</p>
<p>Test security: Measures include video and audio monitoring in test centres and biometrics, including digital photographs and palm vein scanning. Institutions can also access the unscored personal introduction in the speaking section which provides an additional check. Pearson claim to replenish questions continually and randomize test forms to minimize fraud and inappropriate preparation methods. Score reports are only available online through secure logins, as explained above.</p>
<p>Availability and accessibility: The test is available in 186 test centres, including China, India, USA, Japan, South Korea, Australia, the UK, Hong Kong, Taiwan and Canada. Although there are forms available for students requiring scribes or practical assistance there is no obvious link</p>

for students who may have a disability.

Test preparation and test practice resources :

Scored and unscored online practice tests are available and there is also an Official Guide to Pearson Test of English Academic (with CD-ROM) in paperback.

Skills pod for teachers offers online lesson ideas and *skills pod for test-takers* offers online advice and practice, including advice on using commercially available resources (e.g. the range of advanced learner's dictionaries) as well as Pearson's own resources. A Test Taker Handbook is also downloadable in Chinese, Korean and Japanese as well as English.

Contact details:

Pearson Language Tests

80 Strand,

London WC2R 0RL, UK

Telephone : +44 845 543 0243

Fax : +44 20 7010 6611

Evaluation and Comment

PTE academic is a recently introduced test, so there is limited information at present about its impact and the experiences of stakeholders. Standards of security and reliability and user support are obviously a high priority for the providers. Its strengths as a test for academic purposes include the clearly academic focus of its text base, in terms of texts which display, on the whole, academic style and vocabulary. This focus is obtained through the use of their own 37 million word academic corpus. The separation of communicative and enabling skills in rating and reporting is also a useful innovation, as it gives a deeper diagnostic value to the score report, as does the oral personal statement, although that obviously allows for a high degree of rehearsal. The amount of extended writing required is somewhat less than for other major tests and the reading texts are also quite short compared to others.

A major concern about this new test may be the use computer rating for the written and spoken performance as this is a new departure from traditional testing practice. However Pearson claim high correlation between human and machine marking, at 0.96, and that the machine-generated scores explain 92% of the variance of the human ratings. They will continue to rescore randomly selected samples from live administrations of PTE Academic to monitor the accuracy of the automatic ratings. Information about their validity and reliability and the automated scoring procedures as well as other internal research and information about the academic corpus, PICA, are available at <http://pearsonpte.com/research/Pages/home.aspx> This also includes information about the mapping of PTE Academic to the CEFR and alignment of test scores, available at <http://pearsonpte.com/research/Documents/AligningPTEtoCEF.pdf> Details of on-going external research into the PTE academic can be accessed and any completed reports downloaded at: <http://pearsonpte.com/research/Pages/ResearchProjects.aspx>

At the time of writing only one external report is available. However 12 PhD or Masters projects are reported as on-going. Pearson suggest in their document on interpretation of scores that their test discriminates most accurately at levels B1 to C1 of the CEFR descriptors. They point out that cut scores for overall minimum entry requirements should be based on the level and requirements of courses and that minimum sub-scores could be set for certain skills. They point out that the breakdown of enabling skills is less accurate than that of the communicative skills and that these should not be used in high stakes decisions but are useful in diagnosing the type and amount of language support that may be needed. (http://pearsonpte.com/PTEAcademic/scores/Documents/Using_PTE_Academic_v2_13July.pdf)

MELAB - Michigan English Language Assessment Battery **Test Digest**

The MELAB is designed and administered by Cambridge Michigan Language Assessments. It is most widely used in the USA and is rarely presented as a basis for language level evaluation and entry to UK university programmes. However, with the increasing internationalisation of education and education systems and the involvement of Michigan and Cambridge examination bodies with the MELAB test, this may change in the future. MELAB aims to evaluate advanced-level English language competence of adult non-native speakers of English who are applying for academic study to educational institutions where English is the language of instruction. It is also used by professionals who need English for work or training purposes, and those wishing a general assessment of their English level.

Academic Test Characteristics

Skills tested	
Exam length: 4 ½ hours	Mode: Paper-based and audio listening
Use of English and Reading:	
This comprises a GCVR [Grammar, cloze, vocabulary, reading] multiple choice test which incorporates grammar and vocabulary items, reading comprehension tasks, and a 20-item cloze passage, which focuses on the assessment of grammatical structures, vocabulary and reading comprehension. This GCVR assessment has 110 test items. Reading comprises 23 items divided between 4 reading passages.	
The four reading passages deal with four different topics. Reading comprehension items are based on journalistic (e.g. texts <i>National Geographic</i> , <i>New York Times</i> , and <i>Scientific American</i>) in a variety of genres and from topics relating to social sciences, biological sciences, engineering, and business.	
Listening:	
Time: approximately 35 minutes	
This section consists of 60 multiple choice questions, each question having three options. The texts are non-academic and conversational, involving questions, statements, short conversations and radio interviews.	

<p>Writing: Time: 30 minutes Essay; examinees choose one of two prompts designed to elicit expository, argumentative or narrative writing. The writing should consist of 200 – 300 words.</p>
<p>Speaking: There is an OPTIONAL 10-15 minute speaking test. This involves an unscripted one-to-one interview with a certified rater. There are no set questions or prompts, though the interview is divided into three sections, namely a warm-up, extended speaking and a closing segment.</p>
<p>Integration of skills: The different skills are not integrated in the MELAB.</p>
<p>Scores and scoring procedures: Composition Paper: Two trained raters score each composition using a 10-point scale with a third rater scoring in difficult cases. Listening Paper: Scored from between 30 – 100 marks. Scores presented are based on normative information rather than raw scores. GCVR Paper: This is scored from 15-100. As with the listening section, scores presented are based on normative information rather than raw scores. Examinees who do not attempt one of the three sections are given a final score of n.f.s (no final score). Speaking: Possible scores are 1, 1+,2-, up to 4. Key criteria include fluency and intelligibility, interactive ability (including comprehension, conversation development), and language (vocabulary and grammar). Overall: Scores from the composition, listening and GCVR are averaged and an overall score out of 100 is given. Each paper is given equal weight to determine the final score. Students are admitted on to undergraduate and graduate programmes at the University of Michigan with a total score of 80 and no part score less than 80. This is considered to correspond to a score of 84 on iBT TOEFL.</p>
<p>Test security: Examiners are required to keep papers in a secure locked place. Students complete a MELAB identification form on applying for the test which incorporates provision of photo ID, passport or other information for US/Canadian residents. A government issued photo ID must be brought to the examination.</p>
<p>Availability and accessibility: US colleges and test centres in the US and Canada.</p>
<p>Test preparation and test practice resources: <i>A Students Guide to the MELAB 2nd Edition</i> (Spaan, 2008) contains test questions in the MELAB format.</p>
<p>Contact details: Cambridge Michigan Language Assessments Argus 1 Building 535 West William St., Suite 310 Ann Arbor, Michigan 48103-4978 USA Tel: +1 734.615.9629 Fax: +1 734.763.0369 email: info@cambridgemichigan.org</p>

Evaluation and Comment

This test, as far as can be determined, has not been significantly used for entry to universities beyond the USA and Canada, so it is difficult to evaluate this qualification as a basis for academic programme entry in the UK or more generally. The MELAB reading component, while comprising four texts, is limited as a component of the overall test. This is of note as there is some evidence that performance on reading tests is the best predictor in terms of language skills of university performance (for example, Dooley and Oliver, 2002).

The use of the multiple choice GCVR test is inconsistent with the more skills-focused approach of most other tests. Nevertheless, MELAB contains writing, listening and an optional oral test, with the reading assessment embedded as a small component of the GCVR.

With regard to this point, MELAB have commented that research provides evidence that the GCVR measures a single dimension of language ability. MELAB argue that considering reading and cloze items as separate measures from the GCVR is not appropriate. MELAB comment that the GCVR score 'might be interpreted as the examinee's reading score'.

While 30 minutes for the writing test is relatively short compared to the more extended writing found in other tests, the single essay should provide sufficient evidence to judge writing performance. However, the use of a single pre-university essay-genre task may be seen as weakening the tests academic claims.

MELAB test users should take care to pay attention to individual test scores, in particular relating to the composition paper. Low scores here should be seen as of significant concern with regard to university admission. Overall scores may be inflated by high scores in the GCVR test thus which may misleadingly indicate sufficient language ability for university study.

Users in the many courses where speaking is a key component should require a speaking test score in addition to the overall standard MELAB test score.

Annual descriptive statistics and reliability estimates are reported; additionally, a complete MELAB technical manual (2003) contains extensive data regarding the reliability and validity of the test (<http://www.lsa.umich.edu/eli/testing/melab> . Correspondences have been developed at Michigan in relation to the various forms of the TOEFL test⁵.

⁵MELAB/TOEFL CBT comparisons are based on MELAB/Computer-based TOEFL Study, Research Reports 2001-01. English Language Institute, University of Michigan. CBT/PBT comparisons are based on TOEFL Concordance Table, Educational Testing Services, 1998. CBT/iBT comparisons are based on TOEFL iBT/Next Generation TOEFL Score Information, Educational Testing Services, 2004

The TEEP is currently owned and administered by The International Study and Language Centre (ISLC) at the University of Reading as a test for direct entry to UK universities (whether as a stand-alone test or a pre-sessional exit test).

The test originated from an extensive study carried out into the *Language Problems of Overseas Students in Tertiary Education in the UK* (Weir 1983) in response to a growing need to judge the suitability of placing non-native English speaking students on UK university degree courses and was adopted by the AEB (Associated Examining Board, now AQA) during the 1980s. It was redeveloped into its modern format, in 2001 and further modified subsequently. It can be used as a stand-alone UK university entry test; this use expanded gradually both in the UK and overseas until 2011 when UKBA regulations excluded TEEP and any other 'small scale' test from its list of qualifications for obtaining certain types of visa. Although TEEP can at the time of writing be used for direct entry to any UK university which supports its use and will be valid in this situation to obtain a visa, its use is mainly internal at the University of Reading. TEEP is an example of an in-house test which its administrators consider developed, comprehensive, and objective enough to mean that it is not necessary to employ any external assessment on the Pre-sessional English programme.

Academic Test Characteristics

Skills tested Exam length: 3 hours	Mode: Paper based
Language Knowledge Time: 25 minutes 50 questions in 4-option multiple-choice format Mainly grammar, but also related areas such as vocabulary, syntax and linking words. Approximately even distribution of 10 pre-determined language 'areas', attempting to ensure coverage of all relevant grammatical challenges.	
Focus Task: Time: 10 minutes Unassessed, 'brainstorming' exercise. Essay title for the final part is presented, plus space for notes.	
Reading: Time: 35 minutes One passage 1,000-1,200 words. Edited texts from authentic academic sources. General academic style; but not highly technical language. First section is matching heading to paragraphs. Then a series of short-answer questions testing both general and detailed understanding of the text. Candidates are not marked down for spelling or grammar errors (unless serious). Final section is ordering the final paragraph (which is missing from the source).	
Listening: Time: approximately 30 minutes One 'lecture extract', heavily edited from authentic sources, 10-13 minutes long; played	

<p>once. Split into 4 sections. A series of short-answer questions tests both general and detailed understanding of the text. Some gap-filling or multiple choice questions. Candidates are not marked down for spelling or grammar errors (unless serious).</p>
<p>Writing Time: 1 hour An essay on the topic given in the Focus Task, and related to the listening and reading texts. Candidates are expected to use their own ideas as well as ideas retrieved from the reading and listening sections and write in a formal, academic style. All papers are kept until the end, to allow reference to reading and listening sections.</p>
<p>Speaking: Time: 22 minutes for 2 candidates <i>This section was added to the test format in 2011. It is taken separately from the rest of the test.</i> The interlocutor, candidate A and candidate B speak, while the assessor observes. The different parts of the Speaking test are linked to a common topic. There is a file of multiple topics to ensure candidates cannot inform and “prepare” other candidates for the topic. The candidates are given time to read instructions and clarify procedures. The pair discuss a ‘focus question’ to introduce the topic. Each candidate is then given information for a role, supporting one side of the argument. Five bullet points are given to each candidate: each gives a talk (monologues – 3 minutes each) based on these points plus any others they have added. Candidates are then given a scenario with 3 options and discuss possible solutions (dialogue - 4 minutes). In the final stage, the ‘focus question’ is revisited (dialogue – 2 minutes). The interlocutor encourages interaction but tries to stay out of discussions as much as possible. Up to 30% of the 22 minutes is silence, e.g. candidate reading instruction cards and making notes. The examiner gives grades for monologue (global), dialogue (global), plus three analytical criteria: spoken fluency, accuracy and range of grammar and vocabulary, pronunciation/intelligibility. The interlocutor manages the test but also makes global assessments (not analytical).</p>
<p>Integration of skills: The whole basis of the focus task-reading-listening-writing is its topic-linked structure. Reading, Listening and Writing topics are all related. Candidates are expected to use their own ideas as well as ideas retrieved from the reading and listening sections.</p>
<p>Scores and scoring procedures Language Knowledge: Scan-read answer sheet The raw score is converted to one of three ranks ‘below average’, ‘average’, ‘above average’ – based on data collected from current and past administrations – average referring to the “average pre-session English student level.” Reading and Listening: There is double-marking by trained academic staff plus moderation. Keys to short answer questions are by definition open to debate, but are developed with expert agreement and on occasion altered when a suggested change is (near) unanimous. Marks are transferred to an answer sheet, which is then scan-read. Raw scores are converted to 9-band scale, including half-bands – based on trialling results plus past performances (each conversion will be version-dependent). Writing: A standardisation session is held before every administration. There are three criteria: Content/Organisation and Argument/Grammar and Vocabulary – each is scored on a 9-band scale, including half-bands – the overall score is an average of the three. If the two markers’ grades are only 0.5-band apart, the higher grade is used; otherwise an average is</p>

<p>taken. All scripts are double-marked by trained academic staff plus moderation.</p> <p>Speaking: There is a standardisation session held before every administration. There are five criteria: Presenting ideas and information/Interactional skills/Fluency/Accuracy and Range/Intelligibility – each scored on a 9-band scale, including half-bands. The overall score calculation is an average of the five bands, with a slight weighting towards the two global bands.</p> <p>Overall: A 9-band scale, including half-bands. In the case of borderline scores (.25 or .75 calculations), Language Knowledge results act as a decider: 'Above Average' → round up; otherwise round down.</p> <p>The 0-9 scale was brought into use in 2001 and was designed to be 'in line' with IELTS – therefore the default comparison of scores is intended to correspond - something the test providers themselves are not always happy with but the legacy remains.</p>
<p>Test security</p> <p>Every administration is overseen by University of Reading academic staff. All candidates must show passport ID and copies are taken. Invigilation is strict (detailed instructions are given and there are 'floaters' who check everything is running as planned). Certificates are signed in blue, stamped and embossed. Queries direct to the TEEP team are invited.</p>
<p>Availability and accessibility</p> <p>Held at Reading (UK) on 8 dates each year. (Pencil and paper only.)</p> <p>Held in other locations under special arrangement</p>
<p>Test preparation and test practice resources</p> <p>3 practice tests plus advice is available on the website. Teaching course books are available from the University of Reading (these are not designed for self-study). No other publisher currently produces material for the TEEP.</p>
<p>Contact details Bruce Howell / John Slaght International Study and Language Centre, University of Reading Humanities and Social Science (HumSS) Building, Whiteknights PO Box 218, Reading, UK, RG6 6AA Tel: +44 (0)118 378 6477 or 6470 http://www.reading.ac.uk/islc teep@reading.ac.uk</p>

Evaluation and Comment

The test providers' aim is to test academic language and skills as far as is practical in examination format. Conceding that the TEEP does not cover all EAP skills, the test providers justify this on the grounds that no single test can (Howell & Slaght, 2007). The Focus Task and pauses in the Speaking test attempt to provide candidates with time for schemata building. The Reading section time limit encourages expeditious reading, focusing only on the information that is required to complete the tasks. The format of the sources is (quasi-)academic. The essay task is allotted a full hour to encourage planning, synthesising sources with own ideas, and use of referencing. The structure of the linked 'main' test encourages good time management. Both the 'main' test and the Speaking test have a topic theme which is built on, rather than switching topics across tasks and sections.

The Language Knowledge section is justified as being both a 'warmer' (an easy start, in terms of format) and provides useful information for decision-making with borderline students.

All the above points are to be commended – the test providers are attempting to deliver a test which taps into 'EAP' rather than simply English language. The reportedly rigorous marking standards, including regular standardisation sessions are also to be commended. The laborious processes described with invigilating and marking imply that the test is taken very seriously. However, beyond the practice tests and a few now old reports, actual evidence of the quality of the test, statistics showing reliability, error measurement, etc. is lacking. TEEP has a good reputation among a small circle of EAP experts, though this is mainly based on networking, trust and anecdotal rather than hard evidence (though there have been a number of presentations at conferences).

Other criticisms that could be held against TEEP are:

- it may be too topic-dependent – e.g. if a History student meets a topic such as *business practices*, they may perceive that the test is biased against them;
- there is an over-reliance on short-answer questions;
- the once-only listening and the short time for reading limit the range of skills tested;
- it is paper-based, with many pieces of paper and has an 'old-fashioned' feel;
- not enough preparation material is available (for external candidates).

Cambridge Certificate of Proficiency in English

Test Digest

Cambridge CPE covers all four language skills – reading, writing, listening and speaking and a fifth test component 'Use of English' designed to assess such features as morphology, syntax and discourse structure. The test components cover a range of tasks designed to assess the test taker's overall ability to communicate effectively in English and is intended to represent Common European Framework Level C2, learners are expected to be 'approaching the linguistic competence of an educated native speaker' and 'able to cope with high-level academic work' (CPE Handbook).

Academic Test Characteristics

Skills tested:	
Exam length: 5 ½ hours+ 20 minutes speaking based	Mode: Paper
The speaking exam is normally delivered on a separate day*.	
Use of English:	
Time: 90 minutes.	
5 parts with 44 question items. Task types include open cloze, word formation, gapped sentences, sentence transformations, comprehension questions and summary writing task.	

<p>Reading: Time: 75 minutes. 4 passages Approximately 3,000 words overall. The passages include material from newspapers, magazines, journals, books (fiction and non-fiction), promotional and informational materials. 40 questions in total, including multiple choice, gapped text, lexical cloze task types.</p>
<p>Listening: Time: Approximately 40 minutes. 4 parts including <i>monologues, prompted monologues, interacting speakers:</i> discussions, conversations, radio plays, speeches, talks, lectures, commentaries, documentaries, instructions. 28 questions in total involving multiple choice, sentence completion, matching.</p>
<p>Writing: Time: 2 hours. 2 tasks: a compulsory question in Part 1 and one from a choice of five, in Part 2. 300-350 words each. Task types are drawn from the following: an article, essay, letter, proposal, report, review. Each task has a given purpose and a target reader. Each question on this paper carries equal marks.</p>
<p>Speaking: Time: 19 minutes. 3 parts There are two candidates and two examiners. Task types involve: short exchanges with the interlocutor and with the other candidate; a 2-minute 'long turn' and follow-up discussion; a collaborative task involving the two candidates. Content involves personal views and social and topical issues.</p>
<p>Integration of skills: There is no explicit assessment of integrated skills in the CPE exam.</p>
<p>Scores and scoring procedures: The statements of results includes the grades awarded, a graphical display of the candidate's performance in each paper and a standardised score out of 100, allowing candidates to see exactly how they performed. Set values for each grade, allow comparison across administrations of the examination: Grades: A = 80–100 marks; B = 75–79 marks; C = 60–74 marks; Grade D = 55–59 ; E = 54 marks or below. All five papers are weighted equally and a candidate's overall CPE grade is based on the total score gained in all five papers. It is not necessary to achieve a satisfactory level in all five papers to achieve a pass. Results are reported as 3 pass grades (A, B, C) and 2 fail grades (D, E).</p>
<p>Test security: Exam papers are prepared, printed and despatched under secure conditions. All Cambridge ESOL Authorised Test Centres have to follow a detailed code of practice to ensure high standards of security throughout the testing process, from registration to the recording of results; certificates are printed on security-enhanced paper and include other concealed features to prevent forgery and malpractice. The authenticity of certificates can be checked by using Cambridge</p>

ESOL's free Online Verification Service.
Availability and accessibility: CPE is offered in March, May, June and December. Candidates must enter through a recognised centre. The test is taken in around 90 countries worldwide, with the majority in Europe and South America.
Test preparation and test practice resources: A number of coursebooks and practice materials are available from publishers. Care should be taken to ensure that coursebooks and practice materials selected accurately reflect the content and format of the examination. Past papers and examination reports Cambridge ESOL produces past examination papers, for practice, and examination reports, providing a general view of candidates' performance overall and on each paper, and guidance on the preparation of candidates.
Contact details: ESOLinfo@CambridgeESOL.org

Evaluation and Comment

CPE is a well-respected test of English with high standards of reliability and security. The marking and test-setting procedures are robust and examiners are routinely monitored to ensure reliability. Although around 36% of the candidates for CPE report that they are taking the test 'for study', the focus of the test itself is general in nature. While test takers at this level will have a high level of general language proficiency appropriate for most degree programmes, the CPE does not provide an assessment of specific academic or study skills. This is a well-regarded test for general communicative purposes with extensive research to support its validity, reliability and impact for that stated purpose.

Cambridge ESOL offer the following 'Can Do' statements to indicate the typical abilities of a test taker at this level in a study context:

- Listening and speaking – 'CAN understand colloquial asides and cultural allusions'.
- Reading and Writing – 'CAN access all sources of information quickly and reliably. CAN make accurate and complete notes during the course of a lecture, seminar or tutorial'. (CPE Handbook)

This is a far from comprehensive list of the skills needed to cope with the demands of a degree programme and the research for these claims is based solely on self-report data from test-takers most of whom are not preparing for further study. The CPE does not specifically assess the ability to make notes in a lecture or seminar, so this claim is, arguably, questionable.

The exam is scheduled to be revised in 2013, with the aim of making it more suitable for Higher Education purposes. This includes a compulsory essay based on summarizing and evaluating two reading texts.⁶

⁶ <http://www.cambridgeesol.org/assets/pdf/exams/cpe/cpe-bulletin-4.pdf>

Given the high level of language proficiency required at this level (the CPE is intended to discriminate between candidates at the C1 and C2 levels), it is reasonable to assume that CPE test takers achieving a pass grade will not find language to be a barrier to coping with the demands of most English-medium degree programmes. Thus, it would be unnecessary to set CPE as a minimum entry requirement for universities or programmes.

CAE consists of five separate components covering the four main language skills – reading, writing, listening and speaking and a fifth section, Use of English, which covers various sub-skills including vocabulary, morphology, syntax and discourse structure. At this level learners are expected to be able to use the language with ease and fluency. CAE ‘may be used as proof of the level of language necessary to work at a managerial or professional level or follow a course of academic study at university level’ (CAE Handbook).

Academic Test Characteristics

<p>Skills tested Exam length: 4 ½ hours + 15 minutes for the speaking section Mode: Paper-based The speaking exam is normally delivered on a separate day. or Computer based</p>
<p>Use of English: Time: 1 hour 5 parts with 50 questions. Task types include open and multiple-choice cloze, word formation, gapped sentences, key word transformations.</p>
<p>Reading: Time: 75 minutes 4 passages. Approximately 3,000 words overall. 550–850 words per text. The passages include material from newspapers, magazines, journals, books (fiction and non-fiction), promotional and informational materials. 34 questions in total including multiple choice, gapped text and multiple matching task types.</p>
<p>Listening: Time: Approximately 40 minutes 4 parts including <i>monologues</i>: announcements, radio broadcasts, speeches, talks, lectures, anecdotes, etc. and <i>interacting speakers</i>: radio broadcasts, interviews, discussions, etc. 30 questions in total - multiple choice, sentence completion, multiple matching.</p>
<p>Writing: Time: 90 minutes 2 tasks: Task 1: a compulsory question based on reading input of up to 150 words, 180-220 words. Task 2: one from a choice of five, 220-260 words Task types are drawn from the following: an article, a competition entry, a</p>

contribution to a longer piece, an essay, an information sheet, a letter, a proposal, a report, a review. Each task has a given purpose and a target reader. Each question on this paper carries equal marks.

Speaking:

Time: 15 minutes

4 parts. There are two candidates and two examiners.

Task types involve short exchanges with the interlocutor based on photographic prompts; a 1-minute 'long turn'; a collaborative task, involving the two candidates, based on a set of pictures; a discussion on the same topic between two candidates with interlocutor. Topics are general in nature.

Integration of skills: There is no explicit assessment of integrated skills in the CAE exam, although writing task 1 is based on written input.

Scores and scoring procedures: The statements of results includes the grades awarded, a graphical display of the candidate's performance in each paper and a standardised score out of 100 which allows candidates to see exactly how they performed.

Grades: A = 80–100 marks; B = 75–79 marks; C = 60–74 marks; Grade D = 55–59; E = 54 marks or below.

The overall CAE grade is based on the total score gained by the candidate in all five papers. It is not necessary to achieve a satisfactory level in all five papers in order to pass the examination.

Test security: Exam papers are prepared, printed and despatched under secure conditions. All Cambridge ESOL Authorised Test Centres follow a detailed code of practice which ensures the highest standards of security throughout the testing process, from registration to the recording of results; certificates are printed on security-enhanced paper and include other concealed features to prevent forgery and malpractice. The authenticity of certificates can be checked by using Cambridge ESOL's free Online Verification Service.

Availability and accessibility: The CAE is offered at least once per month from February to December. Some administrations are paper-based and others are computer-based. Candidates must enter through a recognised centre.

Test preparation and test practice resources:

A number of coursebooks and practice materials are available from publishers. Most coursebooks will need to be supplemented; care should be taken to ensure that coursebooks and practice materials selected accurately reflect the content and format of the examination.

Past papers and examination reports

Cambridge ESOL produces past examination papers for practice, and examination reports, providing a general view of candidates' performance overall and on each paper, and guidance on the preparation of candidates.

Contact details: ESOLinfo@CambridgeESOL.org

Evaluation and Comment

CAE is a well-respected test of English with high standards of reliability and security. The marking and test-setting procedures are robust and examiners are routinely monitored to ensure reliability. Although around 24% of candidates take CAE 'for

further study', the focus of the exam is itself is general in nature. Cambridge ESOL offer the following 'Can Do' statements to indicate the typical abilities of a test taker at this level in a study context:

- Speaking and listening, 'Can follow up questions by probing for more detail. CAN make critical remarks/express disagreement without causing offence';
- Reading and writing 'CAN scan texts for relevant information and grasp main topic of text. CAN write a piece of work whose message can be followed throughout'.

The research for these claims is based solely on self-report data from test-takers most of whom are not preparing for further study. While test takers achieving a good pass grade at CAE are likely to have a level of language proficiency appropriate for most degree programmes, the test does not assess specific academic or study skills. In particular, the writing tasks bear little resemblance to the kind of writing tasks students on a degree programme are likely to be required to produce as they are based on short texts of around 200 – 250 words and covering mostly general text types such as letters, proposals, reports and articles. This is a well-regarded test for general communicative purposes with extensive research to support its validity, reliability and impact for its stated purpose.

Note that FCE (Cambridge English: First) is sometimes presented for direct entry. However, the focus of the test and its content are general in nature rather than aimed specifically at an academic context; test takers achieving a pass grade at FCE are unlikely to have a level of language proficiency appropriate for most degree programmes. However, FCE may be appropriate for assessing the general English language proficiency of applicant for foundation year programmes or for long term pre-sessional courses.

Business English Certificate Higher (BEC Higher) Test Digest

Developed and administered by University of Cambridge ESOL Examinations (Cambridge ESOL), the BEC Higher examination is aimed at individuals who wish to study a business-related field and at employers who wish to gauge their employees English language level. The content includes business-related topics but does not require specialist knowledge or skills. There are three stages in the examination suite: BEC Preliminary, BEC Vantage and BEC Higher. This digest only covers the BEC Higher.

Test Characteristics

Skills Tested	
Exam Length: 3 hours and 10 minutes + 16 minutes for speaking	Mode: Paper-based
Reading	
Time: 60 minutes	
6 tasks – 52 items	

Task 1: approx. 420 words – identifying specific details – 8 matching items
 Task 2: approx. 330 words – coherence and cohesion – 6 matching items
 Task 3: approx. 580 words – identifying main ideas and details – 6 multiple choice items
 Task 4: approx. 236 words – vocabulary – 10 multiple choice items
 Task 5: approx. 275 words – grammar & vocabulary – five gap-fill items
 Task 6: approx. 180 words – proofreading – 12 items

Writing

Time: 70 minutes – 2 Tasks

Task 1 – Write a report based on graphical information – 120-140 words

Task 2 – Choose between writing a report, a letter or proposal – 200-250 words

Listening

Time: 30 minutes + 10 minutes to transfer answers to mark sheet

3 Tasks – 30 items

Task 1: 2-3 minute monologue – identifying details - 12 fill in the gap items

Task 2: 3 – 4 minutes for five short monologues – listening for gist and details - 10 matching items

Task 3: 4-5 minute conversation or discussion between two or more participants - identifying details - 8 multiple choice items for each task, the recording is played twice.

Speaking

Time: 16 minutes

In the speaking test, two candidates work on three tasks. There are two examiners only one of whom interacts with the candidates.

Task 1: Each individual is asked questions on personal or work-related topics

Task 2: One candidate chooses one topic from a set of three and gives a one minute presentation, the other candidate listens and asks a question at the end. The roles are then reversed.

Task 3: The pair works together to simulate a discussion of a business-related situation.

The candidates are assessed on grammar and vocabulary, discourse management, pronunciation and interactive communication.

Integration of skills: There is no explicit assessment of integrated skills on the BEC Higher exam. However, the writing tasks work from descriptions of business scenarios and brief prompts, including graphs and business letters. The speaking task is also based on written prompts.

Scores and scoring procedures: The statement of results includes the overall grade awarded and a standardised score out of 100. Grades: A = 80–100 marks; B = 75–79 marks; C = 60–74 marks; Grade D = 55–59; E = 54 marks or below. The four skills are weighted equally (25% each) in determining the grade. It is not necessary to achieve a satisfactory level in all four sections to pass the examination. BEC Higher has three passing grades (A, B, & C) and two failing grades (D & E).

Test security: Exam papers are prepared, printed and despatched under secure conditions. Certificates are printed on security-enhanced paper and include other concealed features to prevent forgery and malpractice. Candidates must show a photo ID before exams. Students receive a 'Statement of Results', the authenticity of which can be checked by using Cambridge ESOL's free Online Verification Service.

Availability and accessibility: BEC can be taken as a computer-based examination. Tests are available monthly at centres worldwide. An internet-based examination is not available.

Contact details: <http://www.cambridgeesol.org/exams/index.html> has extensive information

on all aspects of the test.

Test preparation and test practice resources: Cambridge ESOL publishes detailed descriptions of the test types, annual examiners reports and advice to candidates. Several preparation books are available.

Evaluation and Comment

The BEC Higher is one of a suite of three Business English Exams offered by Cambridge ESOL. The BEC exams were originally developed for China at the request of the Chinese National Education Exams Authority in the mid-1990s. The purpose of the tests was to assess the communicative ability of Chinese students who wished to work in international and joint-venture companies. The BEC Higher (previously BEC 3) was last to be developed and added to the suite. In 1998, the BEC exams were made available worldwide and the suite was fully revised in 2002. Despite the claims that the BEC Higher is appropriate for assessing readiness to study business at university level, the primary purpose of the test is to assess Business English as used in professional contexts.

Some of the task types on the BEC Higher mirror task types from other Cambridge ESOL exams which should in theory lead to ease of comparability. However, it is important to note that there are clear differences between the tests which should be taken into account when deciding whether or not to accept scores from the BEC Higher for any particular programme. For example, the overall amount of reading and the length of the individual readings are much shorter in the BEC Higher than for the Cambridge Advanced exam although the two exams are purportedly at the same level of difficulty. The first writing task mirrors Task 1 on the IELTS both in terms of the format of the input and the expected length of the output. However, the second task provides candidates with a choice of writing a report, a letter or a proposal. It could be argued that the formulaic nature of letter writing might advantage candidates who choose that option and also that the task itself is of little relevance in an academic context.

BEC, in common with other Cambridge tests has good standards of reliability and security. The test content aims at a realistic business-related context through tasks such as report writing and business correspondence. It would be useful to see validation studies which demonstrate that the test tasks are equally useful for determining that test takers readiness to perform in academic business environment.

Information on ESP and School level examinations

ESP and Vocational Tests

In some cases HE institutions may consider tests of English designed for specific or vocational purposes for direct entry to academic study into a related subject. Care should be taken that these are designed to give a full assessment of the candidate's English skills. Some may be designed primarily for professional and work purposes rather than for study in that discipline.

ICFE- International certificate for Financial English

This certificate is part of the Cambridge English for Specific Occupational Purposes suite of exams and is intended for use as evidence of the English necessary to work in an international finance context or to follow a course of accountancy or finance at university level. It is aimed to represent levels B2 to C1 of the CEFR. It lasts for 3 ½ hours and tests all four language skills. The reading test takes 1hr 15 minutes and consists of a series of authentic texts from financial professional and academic sources with a total length of about 2,500 words. There are 54 items testing reading skills such as gist, awareness of text structure, implication and detail and also language in use, including business and financial lexis, via cloze tasks and multiple choice.

The listening paper lasts approximately 40 minutes and consists of 4 parts. Part 1 is a series of short extracts from news reports or presentations, with multiple choice questions. Then there is a discussion or interview in a professional financial context with multiple choice questions. Part 4 features a talk on a business topic with gapped phrases to fill in a summary of the talk. Finally there are interpretative matching tasks focusing on aspects of what each speaker says. The four part speaking test lasts 16 minutes, with a paired format and comprises an interview, and extended turn on a finance related topic followed by a collaborative task and a discussion on the same topic.

Results are reported as three passing grades (C1 Pass with Merit, C1 Pass and B2 Pass) and two failing grades (Narrow Fail and Fail). Statements of Results are sent out to all candidates and include a graphical display of the candidate's performance in each paper against the scale Exceptional – Good – Borderline –Weak.

This test provides challenging written and reading tasks which would indicate candidates' ability to cope with academic work in this and related fields. As a pass indicates only overall performance, caution should be exercised in checking the performance in the different skills, paying particular attention to the reading and writing performance.

School Level Examinations

UK-based examinations

These are examinations in English conducted within the UK education system. There are a range of providers in England, Wales and N. Ireland, now under the regulation and oversight of Ofqual (Information available at www.ofqual.gov.uk). In Scotland there is one provider of school and vocational qualifications, the SQA (Information available at www.sqa.org.uk).

The GCSE, IGCSE (First Language Grade C) and Scottish Standard Grade Credit Level and Northern Irish counterparts, Adult ESOL Level 2 and SQA ESOL Higher Level are all accepted as part of the entry requirements for domestic students, so any overseas student presenting with these qualifications at the grade required for domestic students should be accepted on the same terms. Note that the new

English Baccalaureate is not a qualification in itself, but an indication of high performance in a group of core GCSE/IGCSE subjects, including English. In this case it might be advisable to check that the IGCSE English presented is the First Language version (see below).

IGCSE English as a second language is aimed to assess a level of practical communication ideal for everyday use, which can also form the basis for further, more in-depth language study. It is claimed by ICE as suitable for evidence of English proficiency for direct entry at Grade C, but, unlike its First Language counterpart, the texts and tasks on the papers available are very different in level and content from the other tests reviewed and would not prepare students for the demands of academic study. Only reading and writing are tested, and there is no listening. Oral endorsement is also required: syllabus 0511 includes a count-in oral component but 0510 does not. It would be advisable to treat this qualification with caution. It would be probably be suitable only for pre-sessional entry for intending undergraduate students. Further information is available at <http://www.cie.org.uk/qualifications/academic/middlesec/igcse/subjects>

Adult ESOL Level 2 provided as part of the Edexcel skills qualifications suite is also accepted for direct entry for domestic students. The paper-based tests in listening/speaking reading and writing can be taken separately on demand. The reading is assessed through the Adult Literacy National Test. The rather brief tests and the format of a single task for each text make this less challenging than other tests reviewed. However, the preparation for the writing component involves a range of tasks and use of sources, which would be helpful in preparation for academic study. Further information is available at <http://www.edexcel.com/quals/esol/esol-life/levels/Pages/default.aspx>

In considering these school and vocationally based qualifications as evidence of English, it is important to note that these exams are based on descriptors that ensure comparability with other school exams across the curriculum, rather than only levels of language proficiency. This means that credit is given for aspects which would not usually feature in the descriptors of English tests or proficiency frameworks such as the CEFR, for example: standard of cognitive skills, learner autonomy, transferable skills, the complexity demands of the knowledge required and the amount of study involved to achieve them, as well as evidence of study of the course content. These types of requirements might actually be good indicators of capacity for academic study, but will not discriminate specifically between individual students' English proficiency. An illustration of this is that most literate native English speaking adults would be expected to perform at direct entry level in any of the reviewed tests of English proficiency, whereas many native English speakers do not achieve pass grades in the school examinations.

Overseas School Examinations

When deciding on acceptability of overseas school examinations in English, the factors mentioned above should also be taken into account. Such exams may

measure how well the candidates have performed in relation to factors other than simply language proficiency. It is also advisable to compare realistically the likelihood of a UK pupil with GCSE French, for example, being able to study at a French University to the likelihood of a non-native speaker with an equivalent English as a second language school qualification being able to cope with English-medium Higher Education.

Where the exams are taken in the context of English-medium education, it might be expected that the English exam would be evidence of a level of proficiency equivalent to an English GCSE, for example. Although this may be the case, when using such results caution should be exercised on two accounts. One is the extreme variability of what is actually involved in 'English medium education'. Tan and Lan (2011)⁷ report a very varied pattern of delivery with some pupils receiving only key topic words (e.g. the names of chemical elements and compounds) in English, but the lessons being conducted mostly in L1 so that the pupils received no practice in language use. In this case the level of English required in the English exams is likely to reflect this relatively limited exposure to English in the educational environment. In considering school exams, even where A Level or Baccalaureate exams are presented for direct entry, it is advisable to seek evidence on length, content and skills coverage of the exams using the criteria used in the Test Digests. Exams may test writing only in the form of grammar transformations or cloze tests or in essay formats that lend themselves to memorising of large chunks or formats and therefore test only accuracy of reproduction.

A second point in exercising caution is where the form of local English and literacy styles may be radically different from the standard international forms of English and that used in UK academic institutions. University teachers report problems for students for countries such as Nigeria or India, where some students from these areas experience problems in written and oral contexts. It is good policy to encourage such groups of students to attend appropriate level pre-sessional courses to make sure these problems are addressed before they begin academic study.

⁷Tan, M. and Lan, O.S. (2011). Teaching Mathematics and Science in Malaysian classrooms: The impact of teacher beliefs on classroom practices and student learning. *Journal of English for Academic Purposes*, 10, 5-18