

References

- Angoff, W H** (1971) Scales, norms, and equivalent scores, in Thorndike, R L (Ed) *Educational Measurement* (2nd edition), Washington, DC: American Council on Education, 508–560.
- Council of Europe** (2001) *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*, Cambridge: Cambridge University Press.
- Galaczi, E D, French, A, Hubbard, C and Green, A** (2011) Developing assessment scales for large-scale speaking tests: A multiple method approach, *Assessment in Education: Principles, Policy & Practice* 18 (3), 217–237.
- Hawkey, R and Barker, F** (2004) Developing a common scale for the assessment of writing, *Assessing Writing* 9 (2), 122–159.
- Holland, P W and Thayer, D T** (2000) Univariate and bivariate loglinear models for discrete test score distributions, *Journal of Educational and Behavioral Statistics* 25 (2), 133–183.
- Impara, J C and Plake, B S** (1997) An alternative approach to standard setting, *Journal of Educational Measurement* 34 (4), 355–368.
- Jones, N** (2001) The ALTE Can Do project and the role of measurement in constructing a proficiency framework, *Research Notes* 5, 5–8.
- Kolen, M J and Brennan, R L** (2004) *Test equating, scaling, and linking*, New York: Springer.
- Lim, G S** (2012) Developing and Validating a Mark Scheme for Writing, *Research Notes* 49, 6–10.
- Lim, G S, Geranpayeh, A, Khalifa, H and Buckendahl, C W** (2013) Standard setting to an international reference framework: Implications for theory and practice, *International Journal of Testing*.
- Plake, B S and Hambleton, R K** (2000) A standard-setting method designed for complex performance assessments: Categorical assignments of student work, *Educational Assessment* 6 (3), 197–215.
- Shaw, S D and Weir, C J** (2007) *Examining writing: Theory and practice in the assessment of second language writing*, Cambridge: UCLES/Cambridge University Press.
- Taylor, L and Jones, N** (2006) Cambridge ESOL exams and the Common European Framework of Reference (CEFR), *Research Notes* 24, 2–5.
- Jones, N** (2009) A comparative approach to constructing a multilingual proficiency framework: Constraining the role of standard setting, in Figueras, N and Noijs, J (Eds) *Linking to the CEFR levels: Research perspectives*, Arnhem: CITO, 35–43.
- Jones, N and Saville, N** (2009) European language policy: Assessment, learning, and the CEFR, *Annual Review of Applied Linguistics* 29, 51–63.
- Khalifa, H and Weir, C J** (2009) *Examining reading: Theory and practice in the assessment of second language reading*, Cambridge: UCLES/Cambridge University Press.
- Reckase, M D** (2009) Standard setting theory and practice: Issues and difficulties, in Figueras, N and Noijs, J (Eds) *Linking to the CEFR levels: Research perspectives*, Arnhem: CITO, 13–20.
- Taylor, L (Ed)** (2011) *Examining speaking: Theory and practice in the assessment of second language speaking*, Cambridge: UCLES/Cambridge University Press.
- Weir, C J** (2005) Limitations of the common European framework for developing comparable examinations and tests, *Language Testing* 22 (3), 281–300.
- Zeng, L, Kolen, M J, Hanson, B A, Cui, Z and Chien, Y** (2004) RAGE-RGEQUATE computer software, Iowa City, Iowa: CASMA.
- Zieky, M J** (2001) So much has changed: How the setting of cutscores has evolved since the 1980s, in Cizek, G J (Ed) *Setting performance standards: Concepts, methods, and perspectives*, Mahwah, NJ: Lawrence Erlbaum, 19–51.

See also:

Hambleton, R K (2001) Setting performance standards on educational assessments and criteria for evaluating the process, in Cizek, G J (Ed) *Setting performance standards: Concepts, methods, and perspectives*, Mahwah, NJ: Lawrence Erlbaum, 89–116.

Hawkey, R (2009) *Examining FCE and CAE: Key issues and recurring themes in developing the First Certificate in English and Certificate in Advanced English exams*, Cambridge: UCLES/Cambridge University Press.

We are Cambridge Assessment English. Part of the University of Cambridge, we help millions of people learn English and prove their skills to the world.

For us, learning English is more than just exams and grades. It's about having the confidence to communicate and access a lifetime of enriching experiences and opportunities.

With the right support, learning a language is an exhilarating journey. We're with you every step of the way.

Cambridge Assessment English
The Triangle Building
Shaftesbury Road
Cambridge
CB2 8EA
United Kingdom




Copyright © UCLES 2019 | CER/3019/V2/OCT19





**Cambridge Assessment
English**

A decorative graphic consisting of a yellow curved shape on the left, a blue circle in the center, and a pink curved shape on the right, all set against a light gray background.

Comparing scores to IELTS

B2 First and C1 Advanced

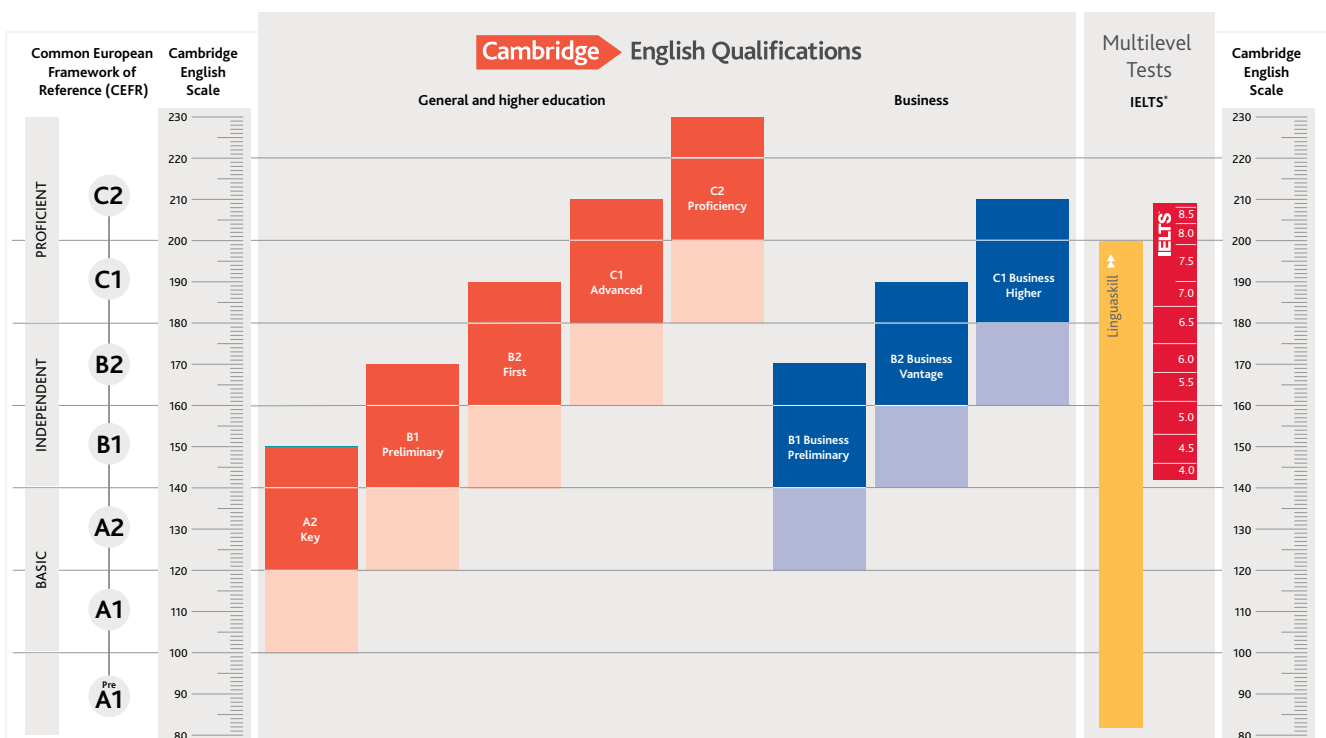
A solid red rectangular block located in the bottom left corner of the page.

Comparing scores to IELTS

B2 First is targeted at Level B2 on the Common European Framework of Reference (CEFR)*. This qualification demonstrates that candidates have the language skills to live and work independently in an English-speaking country or study on courses taught in English at pre-degree level.

C1 Advanced is targeted at Level C1 on the CEFR. It is an in-depth qualification which shows that candidates have the high-level English language skills needed to study in English at undergraduate or postgraduate level, and to work and live in an English-speaking environment.

Each Cambridge English Qualification is focused on a specific CEFR level. For exams from A2 Key to C2 Proficiency, including Business, we also report achievement above and below target level. For Young Learners, we report achievement at the target level and the level below.



B2 First and C1 Advanced both report on the Cambridge English Scale, a single range of scores used to report results for Cambridge English Qualifications, as shown in the diagram above.

Many institutions need to know how Cambridge English Scale scores achieved in B2 First and C1 Advanced compare with IELTS** band scores. The table opposite shows this relationship.

IELTS band score	Cambridge English Scale score
7.5	191
7.0	185
6.5	176
6.0	169
5.5	162
5.0	154

* Further information about the Common European Framework of Reference (CEFR) can be found at cambridgeenglish.org/cefr.

** IELTS is jointly owned by British Council, IDP: IELTS Australia and Cambridge Assessment English.

Using B2 First and C1 Advanced scores

Every successful B2 First and C1 Advanced candidate receives a Statement of Results, which contains the following information:

1. Score – their overall score on the Cambridge English Scale.
2. Candidate profile – their score on the Cambridge English Scale for each of the four skills (reading, writing, listening and speaking) and for Use of English.
3. Grade – relates to the score and is a letter grade from A to C.
4. Candidates also receive an overall level on the CEFR.

Candidates who have secured scores between 160 and 172 on B2 First are awarded grade C on that exam and are placed at Level B2 on the CEFR.

Candidates who have secured a C1 Advanced grade C, having scored between 180 and 192 on the Cambridge English Scale, are at Level C1 of the CEFR and can be expected to be comparable in ability with candidates who have secured 6.5 or 7.0 in IELTS. Candidates who have secured scores between 180 and 190 in B2 First are awarded a grade A for that exam. They are also placed at Level C1 of the CEFR. However, the breadth of coverage of B2 First at this level is limited and very careful consideration would be needed before accepting scores on B2 First as comparable to IELTS scores of 7.0. Candidates who have secured scores of 160 to 179 on C1 Advanced are placed at Level B2 and may be expected to be comparable to candidates who have secured 5.5 or 6.0 in IELTS.

Where institutions have specified a minimum IELTS requirement of 5.5, reference may be made to the Scale score, and a minimum requirement of 162 specified on either exam. If, say, the requirement is Band 6.0 overall but with a minimum score of 5.5 in any skill, then an overall score of 169 may be specified with minimum scores of 162 in Listening, Reading, Writing and Speaking. Where an overall requirement of IELTS Band 7.0 has been set, then a score of 185 should be specified, but as explained above, it may be appropriate to specify that the score has been obtained on C1 Advanced rather than B2 First.

Example requirements

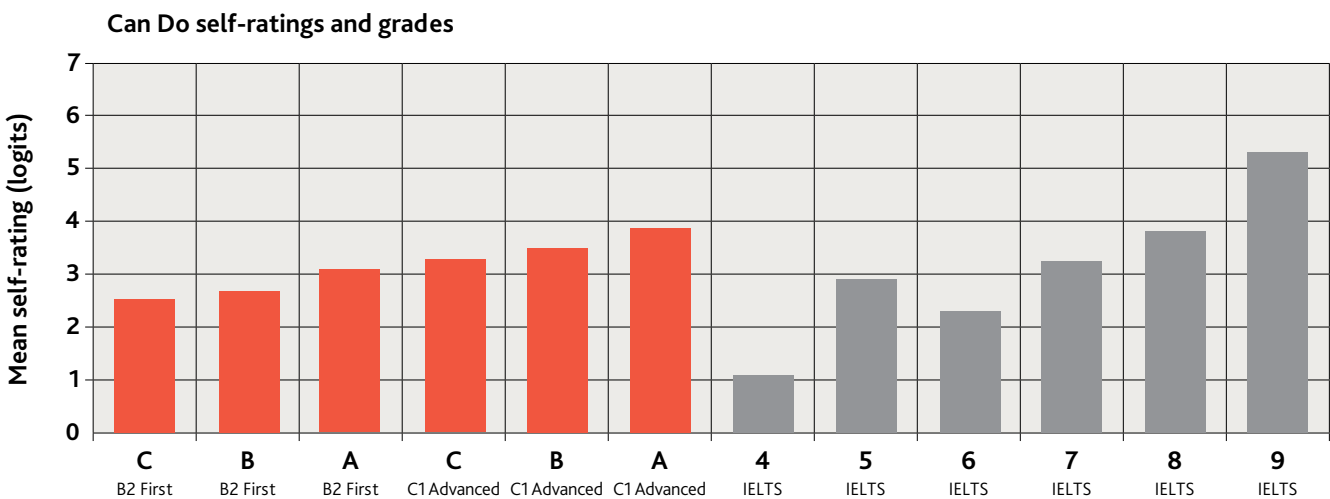
IELTS requirements	Cambridge English requirements
Overall IELTS band score 5.5	Overall Cambridge English Scale score of 162, achievable in B2 First or C1 Advanced.
Overall IELTS band score 6.5 No less than 6.0 in any skill	Overall Cambridge English Scale score of 176 from C1 Advanced. No less than 169 in any paper.

How we compared B2 First scores, C1 Advanced scores and IELTS performances

We are responsible for the production of Cambridge English Qualifications and IELTS. All our qualifications and tests are built to a common underlying scale. Rasch analysis (Rasch 1960, 1980) is used to assess the relative difficulty of every Reading or Listening item, placing each on a common scale, regardless of the exam for which they are intended (Jones 2001). All items are kept in a database with information about their measurement characteristics. This permits the creation of multiple versions of an exam to a specific level and range of difficulty, and establishes the relationship between different exams. We have also published Common Scales for Writing and Speaking, based on qualitative analysis of the features of these skills at different levels (Hawkey and Barker 2004; Galaczi, French, Hubbard and Green 2011; Lim 2012).

Nevertheless, there are challenges associated with linking and comparing exams, as discussed in several Cambridge Assessment English publications (Milanovic 2009; Shaw and Weir 2007; Taylor 2004; Taylor and Jones 2006). Exact equivalences cannot always be demonstrated, only broad comparability. Lim, Geranpayeh, Khalifa and Buckendahl (2013) provide further discussion of the conceptual and practical issues that attend standard setting. It is better not to rely on a single source of evidence but to build up a cumulative case based on a range of data.

Since 1999, several studies have helped refine our understanding of the relationship between these scores. One of the earliest, part of the Association of Language Testers in Europe's Can Do project (Jones 2001), showed that, in terms of candidates' self-perception, candidates securing Band 6 felt comfortable with a similar range of activities as candidates securing a B2 First grade C, while candidates securing Band 7 thought themselves comparable to candidates securing a C1 Advanced grade C. There is a steady progression in self-ratings across IELTS bands (with the exception of Band 5).



In 2009, we undertook to benchmark Level C1 as represented by C1 Advanced against IELTS scores. For this exercise an empirical validation study was undertaken, where registered IELTS candidates were invited to also take C1 Advanced, and registered Advanced candidates were invited to take IELTS, and their scores compared.

This counterbalanced design accounted for preparation or motivation-related effects on one exam or the other. As C1 Advanced targets the higher end of the IELTS candidature population, participants' performance was on average higher than that of the global IELTS candidature, as expected. Correlations between scores on the two exams were calculated to see how related the two tests are. The correlations between the different parts of the two exams are generally moderate, whereas the correlation for the overall scores is, as might be expected, stronger.

To compare results on the two exams, the equipercntile linking method was used, and pre-smoothing using the polynomial log-linear method (Holland and Thayer 2000) was employed to increase the precision of the linking. This method was adopted because indices are available for evaluating goodness of fit and appropriateness of the linking (Kolen and Brennan 2004). Because smoothing resulted in C1 Advanced scores that were not integers, linear interpolation was used to determine IELTS raw marks that corresponded to CEFR Levels B2, C1 and C2 on each of the four skills, and standard conversion tables were used to express the outcomes in terms of the nine-band IELTS scale. Classification consistency between the two exams on the three levels and across the four skills averaged 80%.

In the meantime, the IELTS partners had approached Chad Buckendahl of Alpine Testing Solutions to lead a standard-setting study aligning IELTS bands to the CEFR levels. The standard-setting study involved 19 panellists using two different standard-setting methods for the four papers that comprise IELTS. For Speaking and Writing, a modification of the Analytical Judgment method (Plake and Hambleton 2000) was used. Panellists were asked to read samples of writing and view samples of speaking, and to classify each into appropriate CEFR levels, which was subsequently refined to identify performances at the border of each level. These judgements were then replaced by the original score that those performances received to arrive at the cut score. For Listening and Reading, the Yes/No variation of the Angoff (1971) method (Impara and Plake 1997) was adopted. This standard-setting project is discussed further in Lim, Geranpayeh, Khalifa and Buckendahl (2013).

Our advice as to the alignment of C1 Advanced scores and IELTS bands is therefore based on the results of the external validity study comparing IELTS and C1 Advanced performances, with supporting evidence drawn from the 2009 IELTS standard-setting project and earlier studies.

How was the Cambridge English Scale produced?

There is a well-established link between our qualifications and the CEFR, and the current score-reporting system reflects this. Results on the Cambridge English Scale are reached by applying the same underlying methodology, but the link between our qualifications and the CEFR is refined. This brings a greater clarity and transparency to score meanings and facilitates easy comparisons between different exams.

The Scale was developed according to the well-documented and researched links between performance on different tests (using data from millions of candidates) and the processes by which we define and maintain standards.

These processes vary slightly for the different components and are described below.

Writing and Speaking components

Writing and Speaking components are marked by trained, standardised examiners according to a set of analytic scales, covering a range of assessment criteria. The assessment criteria are linked to the CEFR and form an overlapping 'ladder'. The criteria for each level are the same across all our exams. For example, the criteria required to meet CEFR Level B2 are identical for both B2 First and C1 Advanced.

	B2 First	C1 Advanced	Example assessment criteria
C2			Uses the conventions of the communicative task with sufficient flexibility to communicate complex ideas in an effective way, holding the target reader's attention with ease, fulfilling all communicative purposes.
C1			Uses the conventions of the communicative task effectively to hold the target reader's attention and communicate straightforward and complex ideas as appropriate.
B2			Uses the conventions of the communicative task to hold the target reader's attention and communicate straightforward ideas.
B1			Uses the conventions of the communicative task in generally appropriate ways to communicate straightforward ideas.

Marks are awarded according to the assessment criteria, and are combined to provide the total mark for the component. Because both the assessment criteria and the Cambridge English Scale are linked to the CEFR, the Scale score for the component can be determined from this total mark.

This process ensures that candidates who demonstrate the same level of ability (no matter which exam is taken) are awarded the same Cambridge English Scale score.

For example:

Two candidates at low CEFR Level B2 sit our exams – one sits B2 First, the other C1 Advanced. They both just meet the criteria for Level B2 in the Writing paper and are awarded marks for the component accordingly. Although the raw marks across the two exams are different, the candidates are both awarded a Scale score of 160 for the Writing component, as they have demonstrated the same level of ability.

Reading, Listening and Use of English components

Reading, Listening and Use of English components contain a series of items which are marked as either correct or incorrect. We use Rasch analysis (Rasch 1960, 1980) to ensure a consistent standard is applied in the grading of objectively marked components, accounting for differences in difficulty between them. This is achieved by calibrating the difficulty of all the items in a given test onto the same scale. This calibration allows us to determine the raw marks for each specific test paper that represent a predetermined level of ability – the standard needed to achieve a particular grade or level. Furthermore, the scales used for each test are linked to adjacent levels, meaning that these standards can be compared and linked across levels.

By a process of standard setting, these defined ability levels are linked to CEFR thresholds, meaning that the same process of mapping can take place as with the Writing and Speaking components.

Linking exams to each other and to the CEFR

The relationship between our qualifications and the CEFR is long standing and well documented. The relationship can be classified in terms of the historical perspective, the conceptual perspective and the empirical perspective. Discussions of all three perspectives, plus full references and links to key papers can be found on our website at cambridgeenglish.org/research-and-validation/fitness-for-purpose.

However, test alignment is not a one-off project – validation is an ongoing process which requires regular re-evaluation and confirmation that existing alignments continue to hold.

To this end, and with the introduction of the Cambridge English Scale in mind, a series of alignment studies are in progress to evaluate and validate the links between adjacent exams (for example between B2 First and C1 Advanced) involving candidates taking both exams. This will ensure the integrity of the Scale across our qualifications and tests, and mean that we can be sure that a score of 175 on B2 First corresponds to the same level of performance as a score of 175 on C1 Advanced.

