

Navigating Test Concordance: Supporting meaningful score interpretation in admissions

Leda Lampropoulou, Head of Research

15 October 2026, Higher Education Summit

» The Research Programme: What It's All About

- › Goal: accuracy, fairness, reliability, validity
- › Analysing, calibrating, and validating test materials
- › Monitoring examiners (interlocutors & markers)
- › Assisting in bias and malpractice investigation & prevention
- › Result: Tests maintain the standard & remain fit-for-purpose
- › Evidence published on our website & in peer-reviewed journals

The Science Behind It

- › Classical Test Theory (CTT): reliability, item difficulty, discrimination
- › Rasch: item calibration, test equating
- › Bias detection: Differential Item Functioning (DIF)
- › Marker performance analysis

Research and Validation website

Main strands

- › Innovation and validation studies
- › Calibration studies
- › Empirical studies

- The relationship between language complexity and test-taker achievement on a high-stakes test of writing.
- Delivery of speaking tests in traditional / OLP mode
- Aligning LanguageCert SELT tests to the LanguageCert Item Difficulty scale
- Survey of past OLP candidates' attitudes and perceptions
- Use of pre-task planning time in speaking tests

<https://www.languagecert.org/en/research>



External Validation programme and activities

External validation is achieved through independent audits, reviews and collaborations with HE institutions and their research centres. These have included:

- › ECCTIS – CEFR mapping
- › ALTE's Q-mark (LTE)
- › Concordance study comparing LANGUAGECERT Academic and IELTS Academic (CRELLA)
- › Mapping LANGUAGECERT General to the Canadian Language Benchmarks (CLB) (CRELLA)

> External Validation panel



Prof. Fumiyo
Nakatsuhara

Professor of
Language
Assessment
Director, Centre for
Research in English
Language Learning
and Assessment
(CRELLA)
**University of
Bedfordshire, UK**



Dr Hanan
Khalifa

Expert Member,
**Egyptian
Presidential Council
of Scientists &
Experts, UK**



Prof. Sara
Cushing

Professor of Applied
Linguistics
**Georgia State
University, USA**



Prof. Jim
Tognolini

Director, Centre for
Educational
Measurement and
Assessment
**University of
Sydney, Australia**



Prof. Liying
Cheng

Professor of
Language Education
and Assessment
Director, Assessment
and Evaluation Group
(AEG)
**City University of
Macau**

Concordance Study Design

LANGUAGECERT Academic & IELTS Academic

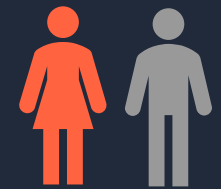


Content comparison

- › Task and item types
- › Scoring rubrics
- › Marking methods

Performance comparison

- › Test performance data
- › Score linking



60% 40%
average age 18-31

1000+

test takers

46

nationalities

~3 mo.

exam
interval

459/549

Counterbalanced
order
(IELTS first/ LCA first)

Official
test reports
only

adhered to all “Good practice principles” (Knoch & Fan, 2024)

Content comparison key findings

Content comparison

The study found a high degree of similarity between the tests' design and the purpose of assessing the ability to use English in an academic setting > it confirmed that the LANGUAGECERT Academic and IELTS Academic tests cover similar content, using similar task types to represent the language needs of students.

Listening

Similar cognitive processes and interaction types
LCA lasts 40 mins including double play vs. IELTS 30 mins single play

Unique to LCA: Listening specifically designed for the Academic domain vs. IELTS used for both the Academic and General modules

Reading

Similar domain: Academic
Similar duration
Similar levels of cognitive processing
Different format and weighting of items

Writing

Same domain: Academic
Similar response formats and genres
Similar tasks: one based on a visual & one essay
Similar expected length of response
Similar weighting
Unique to LCA: performance feedback to support learning

Speaking

Similar duration; one-to-one interview with a human examiner
Unique to LCA: Speaking specifically designed for the academic domain vs. IELTS used for both the Academic and General modules
Performance feedback to support learning
Part 2 role play > interactional competence
Part 4 graph > academic competencies

Scoring comparison key findings

The strong similarities between task and item types, scoring rubrics and marking methods allowed the concordance study to effectively, reliably and meaningfully measure the correlation and equivalences of LANGUAGECERT Academic and IELTS Academic overall test scores and component skill scores.

- › The data collected demonstrate a strong, positive correlation
- › Performance on LANGUAGECERT Academic is **highly predictive** of performance on IELTS Academic

doing well on one test
would translate to doing
well on the other



doing poorly on one test
corresponds with performing
poorly on the other

Correlations					
	Overall r	Reading r	Writing r	Listening r	Speaking r
Academic (n = 1008)	.87	.76	.71	.71	.71
Note: r = Pearson correlation. All correlations were statistically significant at the $p < .001$ level					

Score linking

Methodology: Equipercntile ranking analysis

› While interpreting this table (or any other concordance table), it is important to note that performance on one test - as seen in reported scores - does not guarantee a given performance on the other test.

LANGUAGECERF ACADCMIC					
Overall score	Listening	Reading	Writing	Speaking	IELTS Academic
32	25	26	21	34	4
38	35	36	33	44	4.5
46	41	44	45	54	5
54	49	54	56	62	5.5
61	57	60	64	70	6
67	62	65	71	76	6.5
73	67	71	78	82	7
81	73	77	84	87	7.5
88	80	83	89	89	8
95	89	89	93	93	8.5
N/A	95	97	N/A	99	9

Construct comparability as a prerequisite

Is concordance meaningful? The question starts with constructs, not correlations.

Concordance Tests	TOTAL R	LISTENING	READING	WRITING	SPEAKING
LANGUAGECERT Academic - IELTS Academic	0.87	0.76	0.71	0.71	0.71
IELTS Academic -TOEFL iBT	0.85	0.70	0.76	0.68	0.69
PTE Academic - IELTS Academic	0.78	0.70	0.66	0.56	0.60
MET - IELTS Academic	0.87	0.69	0.75	0.64	0.65
CELPIP General - IELTS Academic	0.84	0.71	0.71	0.68	0.72
C1 ADVANCED - IELTS Academic	0.87	0.78	0.78	0.65	0.77

› The data collected demonstrate a strong, positive correlation

doing well on one test
would translate to doing
well on the other



doing poorly on one test
corresponds with performing
poorly on the other

It does not tell us whether they elicit evidence of language ability in the same way.

Construct comparability as a prerequisite

› Tests differ, sometimes significantly, in the ways information about English language ability is elicited and assessed.

TEST	Writing Task 1	Writing Task 2	Time offered
CAMBRIDGE ADVANCED C1	Essay based on points included in a text	Letter / email, proposal, report or review	90'
CELP	Email	Essay responding to Survey Question	53'
DET	Write a description of an image	Write about a topic and then write a follow-up response	9'
IELTS Academic	Describe a graph, table, chart or diagram	Essay: argument or problem	60'
LANGUAGECER	Academic report in response to an infographic	Essay on an academic subject	50'
PTE Academic	Summarize written text	Write an argumentative essay (& two more tasks)	30'
MET	Respond to personal questions	Write a formal, opinion essay	45'
TOEFL iBT	Write for Academic Discussion	Integrated tasks	30'

Construct comparability as a prerequisite

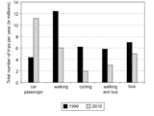
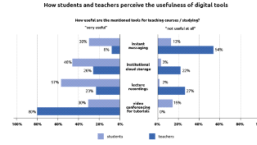
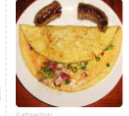
› Tests differ, sometimes significantly, in the ways information about English language ability is elicited and assessed.

TEST	CAMBRIDGE C1 ADVANCED	CELPiP	IELTS Academic	LANGUAGECER Academic	MET	PTE Academic	DET
Task	Write a report / review / letter	Responding to Survey Questions	Write an essay in response to a point of view, argument or problem	Produce a piece of discursive writing on an academic subject	Write a formal, opinion essay	Write an argumentative essay	Write about a topic and then write a follow-up response
Words	220 – 260	150-200	at least 250	250	1-2 pages	200-300	N/A
Time	90' in total	26'	40' / 60'	50' in total	45' in total	20'	5' + 3'

Retrieved from <https://www.cambridgeenglish.org/exams-and-tests/advanced/exam-format/> <https://www.celpip.ca/ta-ke-celpip/test-format/> <https://ielts.org/take-a-test/preparation-resources/sample-test-questions/academic-test> <https://www.languagecert.org/en/preparation/practice-material> https://michiganassessment.org/wp-content/uploads/2020/06/MET-Writing-Resources_FNL.pdf <https://www.pearsonpte.com/pte-academic/test-format/speaking-writing> <https://englishtest.duolingo.com/resources>

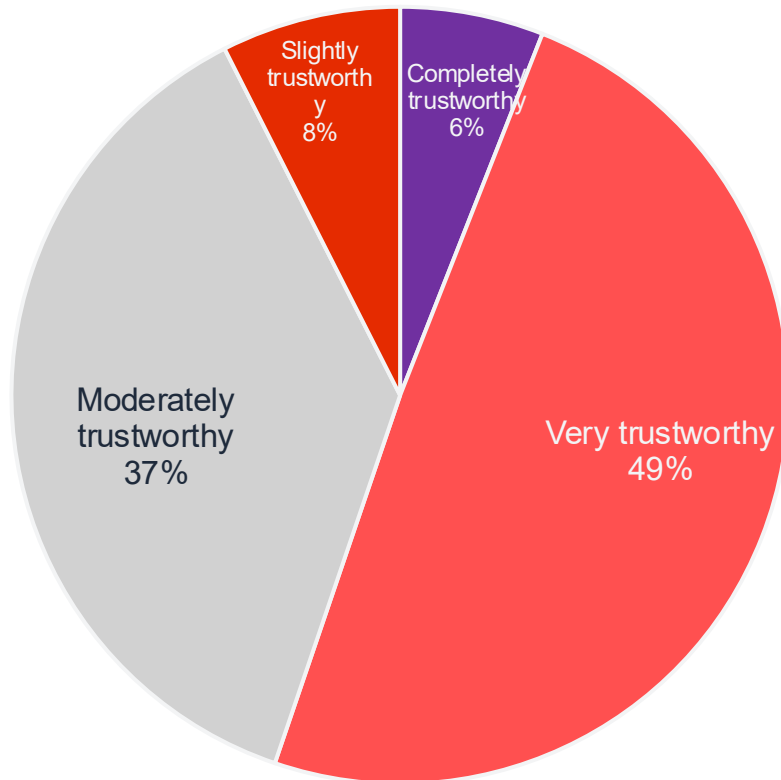
Construct comparability as a prerequisite

Tests differ, sometimes significantly, in the ways information about English language ability is elicited and assessed.

TEST	CAMBRIDGE C1 ADVANCED	CELPIP	IELTS Academic	LANGUAGECERT Academic	MET	PTE Academic	DET
	Write an essay	Write an email	Describe, summarise or explain the information from a graph, table, chart or diagram	Write an academic report or article in response to an infographic stimulus.	Respond to personal questions	Summarize written text	Write a description of an image
Task	Your class has attended a panel discussion on facilities which should receive money from local authorities. You have made the notes below. Which facilities should receive money from local authorities? • museums • sports centres • public gardens Some opinions expressed in the discussion: "Museums aren't popular with everybody" "Sports centres mean healthier people." "A town needs green spaces – parks are great for everybody." Write an essay discussing two of the facilities in your notes. You should explain which facility it is more important for local authorities to give money to, giving reasons in support of your answer.	Read the following information. You recently made reservations for dinner at a very famous and expensive restaurant in town. However, the meal and the service were terrible. The restaurant manager was not available to solve the problem, so you left without a reservation. Write an email to the restaurant's manager in about 150-200 words. Your email should do the following things: • State what problems you had with the food you ordered. • Complain about the service. • Describe how you want the restaurant to resolve the problem to your satisfaction.	The chart below shows the number of trips made by children in one country in 1990 and 2010 to travel to and from school using different modes of transport. Summarise the information by selecting and reporting the main features, and make comparisons where relevant. 	Your university has recently carried out a survey into the attitudes of students and teachers towards various digital tools that are used for either teaching or studying. You have been asked to write a report on the findings of the survey. In your report you should: • Summarise the main findings • Comment on what you think might be the underlying reasons for the findings. How students and teachers perceive the usefulness of digital tools 	* Write sentences to answer the questions. 1. What is your favorite place to travel to? How often have you been there? 2. What do you like about it and why? 3. Tell us about the last time you went there.	Read the passage below and summarise it using one sentence. Type your response in the box at the bottom of the screen. You have 10 minutes to make this task. Your response will be judged on the quality of your writing and on how well your response presents the key points in the passage. For millions of people, smartphones and tablets have been coming to the fore as a means of staying in touch with their family. Family members, from babies born and brought up in the same house to the young adults who have left home, are now able to stay in touch with each other at any time and from anywhere. This has led to a new generation of people who are always connected to their family. This has led to a new generation of people who are always connected to their family. This has led to a new generation of people who are always connected to their family. Write a description of the image below for 1 minute 	
Words	220 – 260	about 150-200	150+	150-200	write sentences	a full, single sentence of no more than 75 words	N/A
Time	90' in total	27'	20' / 60'	50' in total	45' in total	10'	1'

The Trustworthiness Asymmetry

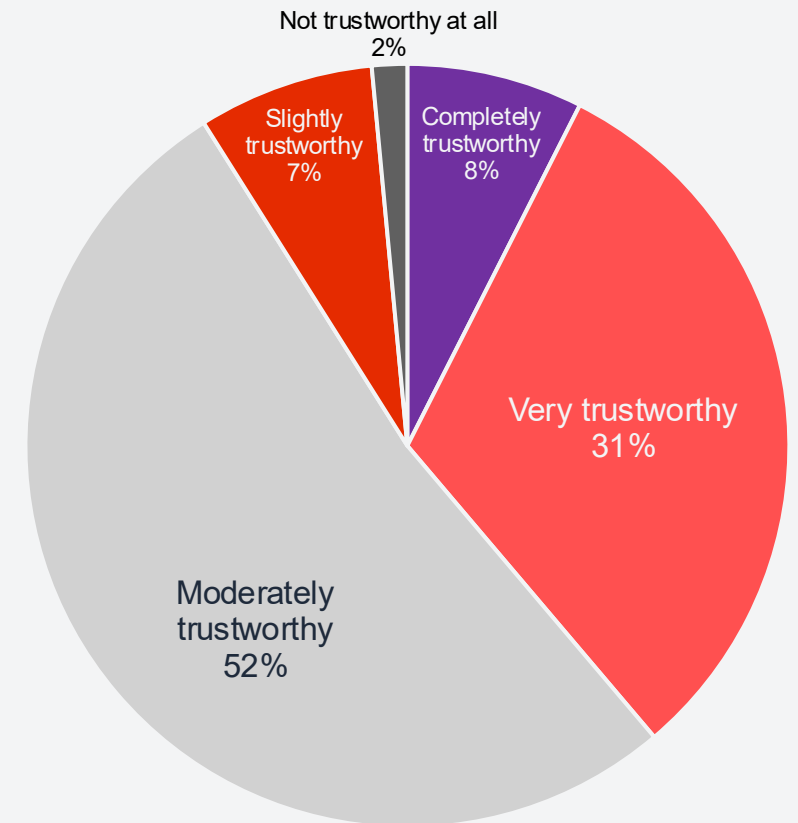
Trustworthiness of test SCORES comparisons



→ Shift to 'moderately' →

- Completely trustworthy
- Very trustworthy
- Moderately trustworthy
- Slightly trustworthy
- Not trustworthy at all

Trustworthiness of test CONTENT comparisons



» Recommendations for interpretation of linkage results

"I am pretty sure [admissions] just take the numbers at face value and don't actually read the studies on which concordance tables are based." *Meg*

"I would recommend making the score comparisons easier to find and understand..."
Peter

"The length of the studies can sometimes be a challenge. Truthfully, we typically skip to the part where equivalent scores are published." *Anne*

"In an ideal world every test would be on a 100-point scale." *Dan*

» Recommendations for interpretation of linkage results

Overall alignment table for LANGUAGECERT Academic and IELTS Academic performance

IELTS Academic	LANGUAGECERT Academic	n-size of study sample at this level	Standard Error
4.5	38-45	27	1.56
5.0	46-53	74	0.79
5.5	54-60	185	0.43
6.0	61-66	225	0.40
6.5	67-72	189	0.46
7.0	73-80	148	0.47
7.5	81-87	104	0.59
8.0	88-94	42	0.78
8.5	95+	3	4.84
9.0	n/a	0	n/a

SE= Standard deviation of LCA scores at each IELTS half band level, divided by the square root of the sample size at that level.

- › Tests often differ in the length of the reporting scales.
- › The choice of concordance study methodology may produce variations in results.
- › The sample sizes used for comparing scores from different tests are generally small across all levels/ranges, especially at extreme ends of the scale.
- › Large Standard Errors show that score equivalences are particularly imprecise at certain points on the ability scale.

› Recommendations for interpretation of linkage results

- › Score comparisons are indicative only and score users are advised not to rely solely on published score equivalences in making their decisions. They should weigh evidence from additional sources where feasible.

LANGUAGECERT ACADEMIC					
Overall score	Listening	Reading	Writing	Speaking	IELTS Academic
32	25	26	21	34	4
38	35	36	33	44	4.5
46	41	44	45	54	5
54	49	54	56	62	5.5
61	57	60	64	70	6
67	62	65	71	76	6.5
73	67	71	78	82	7
81	73	77	84	87	7.5
88	80	83	89	89	8
95	89	89	93	93	8.5
N/A	95	97	N/A	99	9

» Setting minimum score requirements

Understanding the Test Syllabus and CEFR Alignment

Familiarise yourself with the test syllabus, its tasks and skills coverage, and its alignment with the Common European Framework of Reference for Languages (CEFR). This will help you evaluate if the test is useful for the purpose you intend to use it for and help you decide if it can match the necessary language skills for your programmes.

Consultation with Academic Staff

Engage with faculty members and academic staff to map the language skills essential to your programmes, and set cut scores that protect institutional standards and enable student success. Consider your situation.

Benchmarking and Validation

Consider conducting a validation study by tracking the academic performance of students admitted with varying tests / test scores. This will help you refine the cut scores you use over time to ensure they serve your students' and your institution's purposes.

»» Thank you!



Leda Lampropoulou
Leda.Lampropoulou@peoplecert.org

<https://www.languagecert.org/en/research>



by PeopleCert